

QY 360 HisLeuHisCysVal1TriPLeu1LeuAlaArgProGluSerArg1Leu1His1LeuAla1Phe 379
Db 2669 AACATGAACACTGCTGCTGGTATATATCTCGAGACGACAGAACTCGAATTCACCTATATCTTT 2728
QY 380 AsnAsp1LeuAspVal1GluProGlnPheAspPheLeuVal1LeuAspGlyAla1ThrAla 399
Db 2729 AATATATTTGATGTGAGCCCTCAATTTGACTTTCTCCGGCTCAAGATGATGACATTTCT 2788
QY 400 GluAla1ProVal1LeuGly1ThrPheSerGly1AsnGln1LeuProSerSer1Leu1His1Ser 419
Db 2789 GACATTAACGTCTCGGAGTACTTTCTTGCGAATGAGTGCCTTCCAGCTGCGCAGAGT 2848
QY 420 GlyHisVal1AlaArgLeuGluPheGln1ThrAspHis1Ser1ThrGly1AspGly1PheAsn 439
Db 2849 GGGCATATAGTCCCTGCTGGAATTTCACTGTCACATTCACCTACTGCGACAGAGGTTAAAC 2908
QY 440 IleThrPhe1Thr1PheArgHis1AsnGlu1CysProAspProGly1Val1ProVal1AsnGly 459
Db 2909 ATCACTTAACACACNTTTGGTCAGAAATGAGTGCATGATCCTGGCATTTCTATTAACGGA 2968
QY 460 LysArgPheGly1AspSerLeuGln1LeuGly1SerSer1Leu1SerPheLeu1CysAspGluGly 479
Db 2969 CGAGGTTTGGTGAACAGGTTTCTACTGCGAGCTCGGTTCTTCCACTGTGATGATGCG 3028
QY 480 PheLeuGly1ThrGlnGly1SerGlu1Thr1Leu1CysVal1Leu1Gly1SerVal1Val 499
Db 3029 TTTGTCAAGACCCAGGATCGAGTCCATTAACCTGCACTACTGCAAGCGGAAAGTGTGTC 3088
QY 500 TyrAsnSerAlaVal1LeuAsp1CysGluAla1ProCysGly1Gly1His1Leu1His1SerProSer 519
Db 3089 TGGAGCTCCACCGGCTCGCTGGAAGCTCCATGTGTGAGACATCTGACAGCGTCCAGC 3148
QY 520 Gly1Thr1Leu1SerProGly1TyrProGly1Phe1Tyr1Asp1Ala1Leu1SerCysAla1Tyr 539
Db 3149 GGAATCTTTTGGCTCTCGGATGAGCGAGATATTAAGAAATCTTAACTGTGAATGAGG 3208
QY 540 Val1IleGlu1AlaGlnProGly1TyrPro1Leu1Gly1Leu1PheAspArgPhe1His1Gln 559
Db 3209 ATATATGAACCAAAACACGAGCACTGATCAAAATACTTTGACAGATTTCAAGACAGAG 3268
QY 560 Val1Asn1TyrAsp1Thr1LeuGln1Val1AspAspGly1Arg1Thr1Tyr1SerAla1ProLeu1IleGly 579
Db 3269 GTCAATATATACACCTTGAGAGTCAAGATGAGGCGACAGTTCGTCGCCACATGATCGCG 3328
QY 580 Val1Tyr1His1Gly1Thr1Gln1Val1ProGln1Phe1Leu1Ile1Ser1His1Ser1Tyr1Leu1Tyr1Leu 599
Db 3329 GAGTACCAAGGACCCAGGACCCCACTGCTCTCACTCAAGACCGGAACTTCAATGATGACCTG 3388
QY 600 LeuPheSer1ThrAsp1LysSerHis1SerAsp1IleGly1PheGln1Leu1Arg1TyrGln1Thr1Ile 619
Db 3389 CTAATTCACCACTGACAAACAGCCGCTCCAGCATCGGCTTCTCATCCACTATGAGAGTGTG 3448
QY 620 Thr1LeuGln1Ser1AspHis1Cys1LeuAspProGly1Tyr1ProVal1AsnGly1Gln1AspHis1Gly 639
Db 3449 ACCCTTGAGATCGGATTCCTGCTCGAGACCCGCGCATCTCTGTGAACCGCCCATTCGCCAGGT 3508
QY 640 AsnAspPhe1TyrVal1GlyAla1LeuVal1ThrPheSerCysAspSerGly1Tyr1Thr1Leu1Ser 659
Db 3509 GGAGACTTTGGCATCAGGTCCACAGTACATTTTGAAGCTGATCCCGGGGATACACACTAAGT 3568
QY 660 AspArg1Glu1ProLeuGln1CysGlu1ProAsn1PheGln1Tyr1SerAla1Leu1ProSerCys 679
Db 3569 GACGACGAGCCCTCGCTGCTGAGAGGAAACACACAGTGAACACAGCGCTTGCCACAGCTG 3628
QY 680 GluAla1Leu1CysGly1Gly1Phe1IleGln1Gly1Ser1SerGly1Thr1Ile1Leu1SerProGly1Phe 699
Db 3629 GAGCTCTATGTGAGGCTCATCTCAAGGAGAGTGAACAGTCTCTTCTCTCGGATTT 3688
QY 700 ProAspPhe1TyrProAsn1Leu1Asn1Cys1Thr1Tyr1Ile1Gln1Thr1SerHis1Gly1Lys 719
Db 3689 CCAGATTTTATCCAAACTCTTAACCTGACGTTGACCATTTGAAGTGTCTCATGAGGAAA 3748

QY 720 GlyVal1PhePhe1ThrPheHis1SerPheHis1LeuGlu1SerGly1His1Asp1Tyr1Leu1Leu1Ile 739
Db 3749 GAGGTTCAAAATGATCTTTTCAACCTTTCATCTTGAGAGTTCCACAGACATTAATTAAGTATC 3808
QY 740 Thr1Glu1AsnGly1SerPhe1ThrGln1ProLeu1Arg1Gln1Leu1ThrGly1Ser1Arg1Leu1ProAla 759
Db 3809 ACAGAGATGAGAAAGTTTTCGAGCCGCTTGCCAGAGCTCACCGGGCTGGGTGCTTCAT 3868
QY 760 Pro1Leu1SerAla1Gly1Leu1Tyr1Asn1Phe1ThrAla1Glu1Val1ArgPhe1Ile1SerAspPhe 779
Db 3869 ACGATTAAGGACGCGCTGTTGGAAACTTCACTGCGCCACTGCTGTGTTATATAGACTTC 3928
QY 780 SerMet1Ser1TyrGly1Gly1PheAsn1Leu1ThrPheSerGly1Tyr1Asp1LeuGlu1ProCysGlu 799
Db 3929 TCAATTTGATACGAGGCTTCAATATACATTTTCAGAATATGACCTGAGACCATGTGAT 3988
QY 800 Glu1ProGlu1Val1ProAla1Tyr1Ser1Ile1Arg1Gly1LeuGln1PheGly1Val1GlyAsp1Thr 819
Db 3989 GATCTCGAGTCCCTGCTTCCAGCCGAAATGAGTTTCACTTGTGTGGAGACTCT 4048
QY 820 Leu1ThrPhe1SerCysPheProGly1Tyr1Arg1LeuGluGly1Thr1Ala1Arg1Leu1Thr1Ser 839
Db 4049 CTGACGTTTCTCGCTTCTCGGATATCGTTTGAAGGTGCCACCAAGCTTACTCTGCTG 4108
QY 840 Gly1Gly1Arg1Arg1Arg1Leu1Tyr1Phe1Ser1ProLeu1ProArg1CysVal1Ala1Glu1CysGly1Asn 859
Db 4109 GGTGGGGGGCGCGGTGTGTGAGTGCACCTTGCCAAAGTGTGTGGCCAAATGTGAGCA 4168
QY 860 SerVal1Thr1Gly1Thr1GlnGly1Thr1Leu1Leu1Ser1ProAsn1PheProVal1Asn1Tyr1Asn1Asn 879
Db 4169 AGTGTCAAAAGAAATCAAGAAACATTAATCTGTCCAAATTTTCCATCCAAATATATGATAT 4228
QY 880 AsnHis1Glu1Cys1Ile1Tyr1Ser1IleGln1ThrGln1ProGly1Lys1Gly1Tyr1IleGln1Leu1Asp1Ala 899
Db 4229 AACCATGATGTATCTATTAATATGAACAGAAACCGGCAAGGACATTCACCTTAACAAC 4288
QY 900 ArgAla1PheGlu1Leu1SerGlnGly1AspVal1Leu1Lys1Val1Tyr1AspGly1Asn1Asn1Ser 919
Db 4289 CGAGCTTCACAGCTGTTTGAAGAGATCTTAAGATATGATGAGAAAGACACTTC 4348
QY 920 Ala1Arg1Leu1Gly1Val1Phe1SerHis1SerGlu1Met1Met1Gly1Val1Thr1Leu1Asn1Ser1Thr 939
Db 4349 TCACGTCCACTGGGCACTTCACTTAATAATGAATCTTGGGGCTGATCTCTTAACACAGACA 4408
QY 940 Ser1Ser1Ser1Leu1Tyr1Leu1Asp1Phe1Ile1ThrAsp1Ala1Glu1Asn1Thr1Ser1Gly1PheGln 959
Db 4409 TCCAACTACCTGTGGTGAAGTTCACACCAATGATCTGACACCGACCAAGGTTTTCAA 4468
QY 960 LeuHis1Phe1Ser1Ser1PheGlu1Leu1Ile1Lys1Cys1Glu1Asp1ProGly1Thr1Pro1Lys1PheGly 979
Db 4469 CTCACCTATACCAAGTTTATCTGTGTAATATGAGAGATCCGGGCAATCCCTTAACCTAGCGG 4528
QY 980 Tyr1Lys1Val1His1Asp1Glu1Gly1His1Phe1Ala1Gly1Ser1Ser1Val1Ser1Phe1Ser1CysAsp1Pro 999
Db 4529 TATAGATCCGTAAGAAAGCCACTTACCGACACTGTAGTCTGTGACAGTTGCAACCCG 4588
QY 1000 Gly1Tyr1Ser1Ser1Arg1SerGlu1Leu1Leu1Cys1Lys1SerGly1Glu1Tyr1Arg1Thr1Tyr 1019
Db 4589 GGTATGCGCATGATGAGGACGAAACACCTGACCTGTGATGAGAGACAGAGAGTGTG 4648
QY 1020 AspArg1Pro1Leu1Pro1Thr1CysVal1Ala1Glu1CysGly1Gly1Thr1Val1Arg1Gly1Val1Ser 1039
Db 4649 GACAAACCACTACCTTCGTCGATACGGAATGTGGTGGTCAATCATGACACACATCAT 4708
QY 1040 Gly1Gln1Val1Leu1Ser1ProGly1Tyr1Tyr1ProAla1Pro1Tyr1Gln1His1Asn1Leu1Asn1Cys1Ile1Tyr 1059
Db 4709 GGACGAATAATTTTCCCTCGGCTATCCAGTCCGTGTGACAAACACTTCACACGACCTGG 4768
QY 1060 Thr1IleGlu1Ala1Glu1Ala1Gly1Cys1Thr1IleGly1Lys1His1Phe1Leu1Val1PheAsp1Thr1Gln 1079
Db 4769 ATTATTAAGGACAGACCCAGAAAGACATTAAGCTTCATTTATTTTTCACACGAGAG 4828
QY 1080 GluVal1His1AspVal1Leu1Arg1Ile1TyrAsp1Gly1ProVal1Glu1SerGly1Val1Leu1Lys 1099

```
Db 4829 ATGCTCAGACGATCTCTAAGTCTGGGACGGCGCGGTGACAGTGCATCTCTGTAG 4888
Qy 1100 GtLeuSerGlyProAlaLeuProLyAspLeuHisSerThrPheAsnSerValValLeu 1119
Db 4889 GAGTGGAGTGCTCGCCCTTCGCGAGCATCCACAGCACCCTTCATCTCACTCACTG 4948
Qy 1120 GlnPheSerThrAspPhePheThrSerLyGlnGlyPheAlaIleGlnPheSerValSer 1139
Db 4949 CAGTTTCAGACGACGACTTCTTATCATCAGCAAGTCTGCTTCTCATCTCACTTCACTCA 5008
Qy 1140 ThrAlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerGlyAspSer 1159
Db 5009 ATTGCAGCGACCTGTACGATCCAGATGCCCCAAATGCGCACCCTGTATGGAGACAGC 5068
Qy 1160 TrpGluAlaGlyAspSerThrValPheGlnCysAspProGlyIlyTrpAlaLeuGlnGlySer 1179
Db 5069 AGAAGAGCTGGAGACCCGTCACATTCAGATGTACCTGGCTATAGCTCCAAAGACAA 5128
Qy 1180 AlaGluIleSerCysValIleGlnAsnAspPhePheTrpGlnProSerProProThr 1199
Db 5129 GCCAAATATCACCCTGTGACCTGAATATACCGGTTCTTTGGCAACAGACCTCTCTACA 5188
Qy 1200 CysIleAlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeuSerProAsn 1219
Db 5189 TGCAATAGCTGCTTGTGAGGAAATCTACAGGCGCCACAGAGGTATTATTTGTACCCAC 5248
Qy 1220 TyrProGluProTyrProProGlyIlyAspGlnCysAspTrpIlyValThrValSerProAsp 1239
Db 5249 TACCACAGCCGATCTCTCTGGGAGGAAATGTGATGTAGATGAAAGTGAACCCGAG 5308
Qy 1240 TyrValIleAlaLeuValPheAsnIlePheAsnLeuGluProGlyIlyTrpAspPheLeuHis 1259
Db 5309 TTGTTCATCGCTTGATATATCAAAAGTTTCAACATGAGCGCCACCTAGCTTCTCTAC 5368
Qy 1260 IleTyrAspGlyArgAspSerLeuSerProLeuIleGlySerPheTyrGlySerGlnLeu 1279
Db 5369 ATCTATGAGGGAGGAAATCCAAACAGCCCTCATTTGGAGATTACAGAGGCTCTCAGGCC 5428
Qy 1280 ProGlyArgIleGlySerSerSerAsnSerLeuPheLeuAlaPheArgSerAspAlaSer 1299
Db 5429 CCAGAAAGATATAGAGATGAGCGGAAACAGCTGTTTCTGGCATTTGGAGATGATGCTCC 5488
Qy 1300 ValSerAsnAlaGlyPheValIleAspTyrThrGluAsnProAlaGlyIlySerCysPheAsp 1319
Db 5489 GTGGGCTTTTCAGGAGTTCGCAATTAATTAAGAAACACAGCGGAAAGCTGTGTTTAC 5548
Qy 1320 ProGlySerIleLyAsnGlyThrArgValGlySerAspLeuLyLeuGlySerSerVal 1339
Db 5549 CCAGAAATATATGATGAGATGGACAAAGATTGGAAACAGCTTCAGAGCTTGCTTCAACCTC 5608
Qy 1340 ThrTyrTyrCysHisGlyGlyIlyTrpGluValGlnGlyThrSerThrLeuSerCysIleLeu 1359
Db 5609 ACCATACAGTGTGACTGTGCTATATAGATTTCTTGACCCCTCATCATCACTGTGTGATT 5668
Qy 1360 GlyProAspGlyIlyProValIlyPheAsnAspProArgProValCysThrAlaProCysGly 1379
Db 5669 GGGGCTATGGGAAACCTCTCTGGGACCAAGTGTGCTCTCTCATGTGCTCTCTCTGGA 5728
Qy 1380 GlyGlnTyrValGlySerAspGlyValValLeuSerProAsnTyrProGlnAsnTyrThr 1399
Db 5729 GGGCAGTACAGCGGATCCAGAGGGATGTTTATACCAAACTACCCCATATATATACCA 5788
Qy 1400 SerGlyGlnIleCysLeuTyrPheValThrValProLyAspTyrValValPheGlyGln 1419
Db 5789 GGTGTCAATATATGCTCTATTCATCAGTACCGATCAAAAGAAATCGTGTCTTTGGACAG 5848
Qy 1420 PheAlaPhePheHisThrAlaLeuAsnAspValValGlnValHisAspGlyHisSerGln 1439
Db 5849 TTTCCTTAATTTCCAGACAGCCCTGAAATGATTGGCAGAAATTAATTGATGAAACCATGCA 5908
Qy 1440 HisSerArgLeuLeuSerSerLeuSerGlySerHisThrGlyIlySerLeuProLeuAla 1459
Db 5909 CAGGACAGACTTCTCAGCTCACTCTCGGGGTCTCACTCAGGGGAAACATTGCGCTGACT 5968
Qy 1460 ThrSerAsnGlnValIleuIleLyPheSerAlaIlyGlyLeuAlaProAlaArgGlyPhe 1479
Db 5969 ACGCAATATCAAAATTCGCTCCGATTCAGTCAAGGAGCGGTGCTCTGCCCCGGCTTC 6028
Qy 1480 HisPheValTyrGlnAlaValProArgThrSerAlaThrGlnCysSerSerValProGlu 1499
Db 6029 CACTTGTGATTCAGCTGTTCTCTGACAGTACACCCAAATCAGCTCTGTCCCCGAG 6088
Qy 1500 ProArgTyrGlyIlyAspArgLeuGlySerAspPheSerValGlyAlaIleValArgPheGlu 1519
Db 6089 CCCAGATACCGGAAGGAATATGGTTGTGAGTTTCTGCGGGCTCATGTCATTCATTCGAG 6148
Qy 1520 CysAsnSerGlyTyrAlaLeuGlnGlySerProGluIleGluCysLeuProValProGly 1539
Db 6149 TGCAACCCGGGATCTCTGCTTCAAGGTTTCCACGGCGCTCCACCTGCGCTGCTCCAC 6208
Qy 1540 AlaLeuAlaGlnTrpAsnValSerAlaProThrCysValValProCysGlyGlyAsnLeu 1559
Db 6209 GCTTGGACAGTGGAAACGACAGATCCCAAGCTGTGTGATCCCTGACGTGCAATTTTC 6268
Qy 1560 ThrGluArgArgIlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeu 1579
Db 6269 ACTCAAGAGAGGTACATCTCTGCTCCCGGCTACCTGAGCATACGGAACAATTC 6328
Qy 1580 AsnCysValIlyTrpIlyIleValValProGluGlyAlaGlyIlyGlnIleGlnValIlySer 1599
Db 6329 AACTGTATATGAAATCATATGTTACGAGGCGCTCGGAAATTCAGATCCAAATGATAGT 6388
Qy 1600 PheValThrGluGlnAsnTrpAspSerLeuGlnValPheAspGlyAlaAspAsnThrVal 1619
Db 6389 TTTCACAGGAGCAACTGGAGTCCCTTGAGATCCACGATGTGGAGATGATGACGCA 6448
Qy 1620 ThrMetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerThrSerAsn 1639
Db 6449 CCCAGACTGGGAGACTTCTACAGGACCAAGTACCGGCACTGCTGAACAGTATTCAC 6508
Qy 1640 GlnLeuTyrIleuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHisLeuGlu 1659
Db 6509 CAACCTTACCTGCACTTTCAGCTGTGACATTAAGTGTGACAGTGTGCTTCCACTGAA 6568
Qy 1660 TyrIleThrValGlyLeuSerSerCysProGluProAlaValProSerAsnGlyValIly 1679
Db 6569 TACAAACTGTAGTCTTGCTGATCGCAAGACCAAGACCCCTCCACAGCAAGCATCAA 6628
Qy 1680 ThrGlyGluArgTyrLeuValAsnAspValValSerPheGlnCysGluProGlyTyrAla 1699
Db 6629 ATCGGAATGATGATCATAGTGAACGATGCTCTCTTCCAGTCCGAGCCCGGTACACC 6688
Qy 1700 LeuGlnIlyHisAlaHisIleSerCysMetProGlyThrValArgArgTrpAsnTyrPro 1719
Db 6689 CTGCAAGGCGGTCCCAATTTCTGTATGACAGGAGCCGTGCCCTGGAACTATCCG 6748
Qy 1720 ProProLeuCysAlaIleGlnCysGlyGlyThrValIleGlnIleGlnIleGlnValIle 1739
Db 6749 TCTCCCTGTGCAATGTGAACCTGTGGAGGACGTGTGAGCACTTGGGTGTGTATCTGT 6808
Qy 1740 SerProGlyPheProGlyAsnTyrProSerAsnMetAspCysSerTrpIlyIleAlaLeu 1759
Db 6809 AGCCCCGGCTTCCAGTCTTACCCCAACAACTTAACCTGACCTGGAGATCTCATTA 6868
Qy 1760 ProValIlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGluProAsnHisAsp 1779
Db 6869 CCCATCGGTATAGTGCATATATTCAGTTCTGAAATTTTCTACGGAAGCAATATCATGAC 6928
Qy 1780 TyrIleGluIleArgAsnGlyProTyrGluThrSerArgMetCysArgPheSerGly 1799
Db 6929 TTCTTAATTTCAAAATGACCTTACCAACAGCCCAAGATGGAATTAATTAAGCGC 6988
Qy 1800 SerGluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTyrPheHisSer 1819
Db 6989 ACGGATCTCCCGGGCGCTGTGAGCAACGATGAACCCCTCATTCACATTTTATAGT 7048
```

```

QY 1820 Asph1SerGlnAsnArgProGlyPheLeuLeuGluValGlnLeuGlu 1839
DB 7049 GACCATTCGCAAAACCGGACAGGATTTAACTTGCTTACCAAGCTTATGATTTACAGAAC 7108
QY 1840 CysProAspProGluProPheLeuAsnGlyLeuValArgGlyTyrAsnValGly 1859
DB 7109 TGTCCAGATCCACCCCATTTCCAGATGGGTACATGATCACTGATTAAGGATGAGG 7168
QY 1860 GlnSerValThrPheGluCysLeuProGlyTyrGlnLeuThrGlyHisProValLeuThr 1879
DB 7169 CATCAGATCTTCTGAGTGTATTCCTGGGTACATTTAATAGGCACTCTGCTCACT 7228
QY 1880 CysGlnHisGlyThrAsnArgAsnTyrAspHisProLeuProLysCysGluValProCys 1899
DB 7229 TGTCCAGATGGGATCAACAGAACTGGAACCTACCTTTCCAAATGATGATGATGATGAT 7288
QY 1900 GlyGlyAsnLeuIleThrSerSerAsnGlyThrValTyrSerProGlyPheProSerProTyr 1919
DB 7289 GGGTACACAGTACCTTCTCCAGACGCGCACATCTACTCCCTGGCTTCCGATGATGAT 7348
QY 1920 SerSerSerGlnAspCysValThrPheLeuIleThrValProLeuGlyHisGlyValArgLeu 1939
DB 7349 CCGATCCTGGAAGGACTGATTTGGCTCATTCAGGAGCTCCAGGCGACGAGTTACATC 7408
QY 1940 AsnLeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTyrAspGlyPro 1959
DB 7409 AACTTCACCTGTTACAGACGGAAGCTGTCAACGATTAACATTTGCTTGGAGCGGCC 7468
QY 1960 GlnGlnThrAlaProAlaGluGlyValPheThrArgSerMetAlaLysValThrValGln 1979
DB 7469 GATCAAGACTCAGCCAGCTGGAGTTTCAGTGGACACACAGCCCTCCAGCGGTAT 7528
QY 1980 SerSerSerAsnGlnValLeuLeuLysPheHisArgAspAlaIleThrGlyGlyIlePhe 1999
DB 7529 AGCTCCACCAACCAAGCTCTGCTCAAGTTCCACAGCAGCTTTCAAAATGAGGCTTTT 7588
QY 2000 AlaIleAlaPheSerAlaTyr-----ProLeuThrLys----- 2010
DB 7589 GTCTCAATTTCACGCTGATGATTTTCATCCGTTAGTAAAGTGAATTCAGT 7648
QY 2011 -----CysProPro-----ThrIleLeuProAsnAla 2020
DB 7649 TGGTGTACTGCACTGTGTTCACGCTTTTCCAGCTGAAGTTTGTGATCAGCC 7708
QY 2021 GluValValThrGluAsn 2026
DB 7709 GAGGCGGTGATGATCT 7726

```

RESULT 6
US-10-276-934-3
Sequence 3, Application US/10276934
Publication No. US20030180750A1
GENERAL INFORMATION:
APPLICANT: University of Leeds
APPLICANT: Markham, Alexander F.
APPLICANT: Jackson, Andrew P.
APPLICANT: Woods, Christopher G.
FILE OF INVENTION: Treatment of Cancer and Neurological Diseases
TITLE REFERENCE: 9052-144
CURRENT APPLICATION NUMBER: US/10/276,934
PRIOR FILING DATE: 2002-11-20
PRIOR APPLICATION NUMBER: PCT/GB01/02240
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: GB0012186.3
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 6409
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

```

NAME/KEY: misc feature
LOCATION: (588)..(588)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (624)..(624)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (672)..(672)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (2265)..(2265)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (6387)..(6387)
OTHER INFORMATION: "n" is any nucleotide
US-10-276-934-3

Alignment Scores:
Pred. No.: 0
Score: 7610.50
Percent Similarity: 79.70%
Best Local Similarity: 65.74%
Query Match: 44.81%
DB: 15 Gaps: 10

US-10-016-248-2 (1-3104) x US-10-276-934-3 (1-6409)
QY 36 ValLysLysGlnIleGluLeuLysSerArgValLysLeuMetProSerLysAspAsn 55
DB 66 GCGAAAGAGGATTTGATGAGTCAAGAGATCAAGATCTGCCAGCAAGATGCA 125
QY 56 SerGlnLysThrSerVal----- 61
DB 126 ACCCATTAATACTGTCTGGCATATGAAATACATCAACCATGATCCGAGAGCA 185
QY 62 -----LeuThrGln-ValGlyValSerGlnI 70
DB 186 TACGTTCAATTATGATCTGGGAAACCATGCTTCACAAATTCAGATGACCAAG 245
QY 70 Y-----HisAsnMetCysProAspProGlyIleProGlyValArg 85
DB 246 AGCTGTGATTTGCTGCTGCTGATGTCTCCAGATCTGGATTCAGAAATGATGAG 305
QY 85 GlnGlySerAspPheArgLeuGlySerSerValGlnPheThrCysAsnGlyTyrAs 105
DB 306 AGCAGGTTCCGACTTCAGGTTGGTGCAATGTACATGATTCATGATGAGCAATTA 365
QY 105 PLeuGlnLysSerLysArgIleThrCysMetLysValSerAspMetPheAlaIleTyr 125
DB 366 GCTCCAGGATCTTAAGACATCACTGTCAGAGGATTAACAAGCTGCTGCTTGA 425
QY 125 AspHisArgProValCysArgAlaArgMetCysAspAlaHisIleuArgGlyProSer 145
DB 426 TGACCAAGCGCCCATCTCCAGAGGAAACATGTGATCCAAATCTGCTGGCCAGCG 485
QY 145 YIleIleThrSerProAsnPheProIleGlnTyrAspAsnAlaHisCysValTyr 165
DB 486 COTCATTTACCTCCCTTAATATCCGTTCCAGTATGAGATATACACACTGTGTGG 545
QY 165 eIleThrAlaLeuAsnProSerLysValIleLysLeuAlaPheGlnGluPheAspLeu 185
DB 546 CATCCACCAACCGACCGGACAAAGTCAATCACTTCCCTTNAAGAATTTGAGCTGA 605
QY 185 ValArgGlyTyrAspThrLeuThrValGlyAspGlyGlyGlnAspGlyAspGlnLys 205
DB 606 GCGAGGCTATGACACCTTACCGTTGATGATGCTGGGAGAGTGGGACACCAATGCT 665
QY 205 IleuTyrMetSerGlnAsnAlaCysSerAspSerProHisThrProGlySerArgIle 225

```


Db 666 CTGTGATG-----CTCAGGGATCCAGTGTCC 695
 Qy 225 oglu-----serMetSerGlyAspIleThrArgInlyStrpThrValLeuGluI 242
 Db 696 TGACCTCATTTGTGAGCATGACCAACCAAGATGTGCTACAT----- 735
 Qy 242 eCyArgAspIleSerSerSerAspAlaArgSerGlySerValArglySerProlyeTh 262
 Db 736 -----CTCAGATCGGATGATAGC----- 753
 Qy 262 rSerAsnAlaValGluLeuValAlaProGly-----ThrGluIleGluI 277
 Db 754 -----ATTGCTCAGCTGGGTTTAAAGCTGTATTACCAAGAAATTGAAA 797
 Qy 277 nglySerCyArgIlyAspProGlyIleProAlaTyGlyArgArgGluGlySerArgPheH 297
 Db 798 GGGAGGGGTGGGATCTCTGGAAATCCCGCTATGGGAAGGAGGAGGAGGAGGAGTTCCT 857
 Qy 297 bHlGlyAspThrLeuLysPheGluCyArgInProAlaPheGluLeuValGlyInlybAl 317
 Db 858 CCATGGAGATACACTCACCCTTTGAATGCCGGGCTTTTGAAGCTGGTGGGAGAGAGT 917
 Qy 317 aIleThrCyArgInlybAsnAngInlyrSerAlaLysLysbProGlyCybValPheSerCy 337
 Db 918 TATCACCCTGCAGAGAACATCAGTGTGGCAACAGCCAGCTGTATTTCATG 977
 Qy 337 bPhePheAsnPheThrSerProSerGlyValAlaLeuSerProAsnTyTProGluAspTy 357
 Db 978 TTTTTCACATTTACGGCATCATCTGGGATTAATCTGTCAACAAATTATCCAGAGGATATA 1037
 Qy 357 rGlyAsnHlAsnLysCybValTProLleLeuAlaArgProGluSerArgIleHlSle 377
 Db 1038 TGGGAACAACATGAACTGTGTCTGTGATTATCTCGAGCCAGAGAACTCAATTACCT 1097
 Qy 377 uAlaPheAsnAspIleAspValGluProGlnPheAspPheLeuValIleLysAspGlyAl 397
 Db 1098 AATCTTATATGATTTTGAATGTTGAGCCATTTGATTTCTCGGGGTCAAGATGATG 1157
 Qy 397 aThrAlaGluAlaProValLeuGlyThrPheSerGlyAsnGlnLeuProSerSerIleTh 417
 Db 1158 CATTCTGCACATTAACCTGTCCGGGTACTTTTCTGGCAATGAAAGCTTCCCAAGCTGGC 1217
 Qy 417 rSerSerGlyHlSleValAlaArgLeuGluPheGlnThrAspHlSerThrGlyLysArgG 437
 Db 1218 CAGAGATGGGCATATAGTTCCTTGAATTCAGATTCAGATTCCTACTGCGAGAGG 1277
 Qy 437 yPheAsnIleThrPheThrThrPheArgHlAsnGlnCyAspProAspProGlyValProVa 457
 Db 1278 GTTCAACATCACTTACACCAATTTGTGCAGATGAGTGCATATCTGSCATTTCTCAT 1337
 Qy 457 lAsnGlyLysArgPheGlyAspSerLeuGlnLeuGlySerSerIleSerPheLeuCybAs 477
 Db 1338 AAAGCGAGCAGTTCCTGTCAGAGTTCCTACTCGGGAGCTCGGTTCTTCCACTGTGA 1397
 Qy 477 pGluGlyPheLeuGlyThrGlnGlySerGluThrIleThrCybValLeuLysGluGlySe 497
 Db 1398 TGAATGGCTTTGTCAAGACCCAGAGTCCGAGTCCATTAACCTGCACTGCAAGAGGGAA 1457
 Qy 497 rValValTPraSerAlaValAlaLeuArgCyArgIlyAlaProCyArgIlyHlSleLeuThSe 517
 Db 1458 CGTGGTCTGAGCTCCACCGTCCCGCTGTGAAGCTCATGTGGTGAACATCTGACAGC 1517
 Qy 517 rProSerGlyThrIleLeuSerProGlyTPProGlyPheTyryLysAspAlaLeuSerCy 537
 Db 1518 GTCCAGCGAGTCAATTTTGTCTCTCGATGCGCAGATATATTAAGATTCCTTTTAACTTG 1577
 Qy 537 bAlaTPraValIleGluAlaGlnProGlyTyTProIleLysIleThrPheAspArgPheTy 557
 Db 1578 TGAATGGATTAATGAAGCAAAACAGGCACTCTATCAAAATTAATCTTTGACAGATTTCA 1637
 Qy 557 bThrGluValAlaAsnTyArgbThrLeuGluValArgAspGlyArgThrTyrSerAlaProLe 577
 Db 1638 GACAGAGGTCAATTAATGACACTTGTGAGGTCAAGATGCGCCAGCCAGCTTGTCCCACT 1697

Qy 577 uIleGlyValTyThrHlGlyThrGlnValProGlnPheLeuIleSerThrSerAsnTyLe 597
 Db 1698 GATCGGAGAGTACCCAGGACCCAGGACCCCACTCTCATTCAGCACCGGGAATTCAT 1757
 Qy 597 uTyrLeuLeuPheSerThrAspLysSerHlSleAspIleGlyPheGlnLeuArgTyG 617
 Db 1758 GTACTCTCTATTACCACTGACCAACACCGGTCAGCATGGCTTCCATCTCACTATGA 1817
 Qy 617 uThrIleThrLeuGlnSerAspHlSleAspProGlyIleProValAsnGlyInAr 637
 Db 1818 GAGGTGACCGCTTGAAGTCGATTCCTCGAGCCCGGAGATCCCTGTGAACGGCCATCG 1877
 Qy 637 bHlGlyAsnAspPheTyrValGlyAlaLeuValThrPheSerCybAspSerGlyTyTh 657
 Db 1878 CCACGGTGGAGACTTGTGACATCAGGTCAAGTACATTTTCAAGCTGTACCCGGGATCAC 1937
 Qy 657 rLeuSerAspGlyValProLeuGluCybGluProAsnPheGlnTPraSerArgAlaLeuPr 677
 Db 1938 ACTAAGTACAGACAGGCCCTCTGTGTGAGAGGAAACCAAGTGAACCAAGCCCTTGCC 1997
 Qy 677 oSerCyArgIlyAlaLeuCyArgIlyPheIleGlnGlySerSerGlyThrIleLeuSerPr 697
 Db 1998 CAGCTGCAGCGCTTATGTGAGGCTACATCCAAAGGAGAGTGAACAGTCTTCTTCC 2057
 Qy 697 oglyPheProAspPheTyrProAsnLeuAsnCybThrTProIleIleGluThrSerHl 717
 Db 2058 TGGGTTCACAGATTTTATCCAACTCTTAACATGCAACGAGCAATGAAAGTGTCCA 2117
 Qy 717 sGlyLysGlyValPhePheThrPheHlStrPheHlLeuGluSerGlyHlSleAspTyLe 737
 Db 2118 TGGGAAGAAGATTCAATGATCTTTCACACTTTCATCTTGAAGTTCACCAAGCATATT 2177
 Qy 737 uLeuIleThrGluAsnGlySerPheThrGlnProLeuArgGlnLeuThrGlySerArgLe 757
 Db 2178 ACTGATCAGAGAGATGAAGTTTTCGAGCCCGTGTCCAGAGCTCAGCGGTGGTGT 2237
 Qy 757 uProAlaProIleSerAlaGlyLeuTyrglyAsnPheThrAlaGlnValArgPheIleSe 777
 Db 2238 GCTTCATACGATCAAGCAGCGCTGTGGAATTCATCTGACCAGCTTGTGGTTTATATC 2297
 Qy 777 rAspPheSerMetSerTyrglyGluPheAsnIleThrPheSerGlyTyArgbLeuGluPr 797
 Db 2298 AGACTTCTCAATTTTCGACAGGCTTCATATACATTTTCAGAAATATGACCTGAGACC 2357
 Qy 797 oCyArgIlyProGluValProAlaTySerIleArgLysGlyLeuGlnPheGlyValG 817
 Db 2358 ATGTATGATATCTGAGAGTCCCTGCTCAGCCGAAATGTTGATTTCACTTGTGTGGG 2417
 Qy 817 yAspThrLeuThrPheSerCybPheProGlyTyArgLeuGluGlyThrAlaArgIleTh 837
 Db 2418 AGACTCTCTACCGTTTTCCTGCTTCGGAATATCGTTTGAAGATGCCCAAGCAAGCTTAC 2477
 Qy 837 rCybLeuGlyValArgArgArgLeuTPraSerProLeuProArgCybAlaGluCy 857
 Db 2478 CTGCTCGGTGGGGGCGCGCTGTGTGAGTGCACCTCTGCCAAGGTGTGTGGCGGATG 2537
 Qy 857 sGlyAsnSerValThrGlyThrGlnGlyThrLeuLeuSerProAsnPheProValAsnTy 877
 Db 2538 TGAAGCAAGTGTCAAGAAATGAAGAACATTAATCTGTCCAAATTTTCCATCAATTA 2597
 Qy 877 rAsnAsnAsnHlSleGlyValleTySerIleGlnThrGlnProGlyLysGlyIleGlnLe 897
 Db 2598 TGAATTAATACCATAGGTATCTATAAATGAAGAAAGGCGGCAAGGAGCATCCACT 2657
 Qy 897 uTyAlaArgAlaPheGluLeuSerGluGlyAspValLeuLysValTyArgbGlyAsnAs 917
 Db 2658 TAGAACAGAAAGCTTCAGCTGTTTGAAGAGATCTTAAAGATATATGATGAAGAAAGA 2717
 Qy 917 naenSerAlaArgLeuLeuGlyValPheSerHlSleGluMetMetGlyValThrLeuAs 937
 Db 2718 CAGTTCCTCAAGTCACTGGGCACTTCACTAAATAAGAACTTGTGGGCTGATCTCAAA 2777

```

Qy 937 nSerThrsSerSerLeuTrpLeuAspPheIleThrsAspAlaGluAenThrSerLeu 957
Db 2778 CAGCACATCATCATCACTGTGGCTAGAGTTCAACCAATGAGATGTGACACCAAG 2837
Qy 957 yPheGluLeuHisPheSerSerPheGluLeuIleCysGlySerProGlyThrProly 977
Db 2838 TTTTCACTCACTATACCACTGTTTGAATCTGTATAAATGATGAGATCCGGACATCCCTAA 2897
Qy 977 sPheGlyThrLeuValHisAspGluGlyHisPheAlaGlySerSerLeuSerPheSerC 997
Db 2898 CTAGGCTATAGAGATCCGTGATGAAAGGCCATTATACCACTGATGTTCTGTACAGTTG 2957
Qy 997 sAspProGlyThrSerLeuArgGlySerGluLeuLeuCysLeuSerGlyGluArg 1017
Db 2958 CAACCCGGGTTACCGCATGCGACCAACCCCTGACCTGTTGATGAGACAGAGAG 3017
Qy 1017 gThrTrpAspArgProLeuProThrCysValAlaGluCysGlyGlyThrValArgGly 1037
Db 3018 AGTGTGGAGACAAACCACTACCTGCTGATGACGGAATGTGTGTGATCAGATCCATGCGAGC 3077
Qy 1037 uValSerGlyLeuValLeuSerProGlyThrProAlaProTyrGluHisAsnLeuAenC 1057
Db 3078 CACATCAGAGCAGAAATATTGCTCCCTGGCTATCCAGCTCCGATACAAACCTCCACCTG 3137
Qy 1057 sIleTrpThrIleGluAlaGlyCysThrIleGlyLeuHisPheLeuValPheAs 1077
Db 3138 CACCTGATTTATAGGAGGACAGCCAGAAACCATTAAGCTTCATTTGTTTGA 3197
Qy 1077 pThrGluGluValHisAspValLeuArgIleTrpAspGlyProValGluSerGlyValLe 1097
Db 3198 CACGAGATGGCTACACATCTCTCAAGCTTGGAGAGGGCCGGTGGACAGTGCACATCT 3257
Qy 1097 uLeuLeuGluLeuSerGlyProAlaLeuProLyAspLeuHisSerThrPheAsnSerVa 1117
Db 3258 GCTGAAGAGAGTGGAGTGGCTCCGCCCTCCGAGAGATCCACAGCACTTCAACTACT 3317
Qy 1117 lValLeuGlnPheSerThrAspPheThrSerLeuGlnGlyPheAlaIleGlnPheSe 1137
Db 3318 CACCTTCAGATTGACACGCACTTCTTCATCAGACATCTGCTTCCATCCAGTTCTC 3377
Qy 1137 fValSerThrAlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerG 1157
Db 3378 CACCTCAATTGGACACCACTGTATACATCCAGATACCCCAAAATGGACCCGCTATGG 3437
Qy 1157 yAspSerTrpGluAlaGlyAspSerThrValPheGlnCysAspProGlyTyrAlaLeuG 1177
Db 3438 AGACAGAGAGAGGCTGAGACACCTGCATTCAGATGAGACCTGCTATCAGCTCCA 3497
Qy 1177 nGlySerAlaGluLeuSerCysValIleGluAsnArgPheThrGlnProSerP 1197
Db 3498 AGGACAAAGCCAAATACCTGTGTGACGCTGAATTAACCGGTCTTTTGGCAACCAAGCC 3557
Qy 1197 oProThrCysIleAlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeuSe 1217
Db 3558 TCCTCAATGATGATGCTGCTGTGAGAGGATCTGACGGGCCACAGAGTGTATTTTGTTC 3617
Qy 1217 rProAsnTyrProGluProTyrProProGlyGlyGlyCysAspTrpLeuValThrValSe 1237
Db 3618 ACCCAACTACCAACAGCCGATCTCTCTGGAGAGATGATGATGAGAGTAAAGTGA 3677
Qy 1237 rProAspTyrValIleAlaLeuValPheAsnIlePheAsnLeuGlnProGlyTyrTrpAs 1257
Db 3678 CCGGAGCTTTGATGCTGCTGTGATTAATCAAAAGTTTCAACATGAGACCCAGCTATACCT 3737
Qy 1257 eLeuHisIleTyrAspGlyArgAspSerLeuSerProLeuIleGlySerPheTyrGlySe 1277
Db 3738 CTTACACATCTATGAGAGGAGATTCCAACAGCCCTCATTTGGAGATTACAGGGCTC 3797
Qy 1277 rGlnLeuProGlyValGlyIleGlySerSerAsnSerLeuPheLeuAlaPheArgSerAs 1297
Db 3798 TCGAGGCCCGAGAAAGATAGAGTAGCGGAAACACCTGTTTCTGGCATTTCCGAGTGA 3857
Qy 1297 palAspValSerAsnAlaGlyPheValIleAspTyrThrGluAsnProArgGluSerC 1317

```

```

Db 3858 TGCCCTCCGTGGGCTTTCAGAGGTTCCCATTAATTAAGAAACCAAGGAGCTTG 3917
Qy 1317 sPheAspProGlySerIleLeuAsnGlyThrArgValGlySerAspLeuLeuGlySe 1337
Db 3918 TTTTGACCCAGAAATATATGAAATGGAACAGAGATTGGAAACAGATTCAGACTGGCTC 3977
Qy 1337 fSerValThrTyrTyrCysHisGlyGlyThrGluValGluGlyThrSerThrLeuSerC 1357
Db 3978 CACCATCACTACCACTGATGACTCTGTGATTAAGATTTTGAACCCCTCATCACTACCTG 4037
Qy 1357 sIleLeuGlyProAspGlyLeuProValTrpAsnAspProArgProValCysThrAlaP 1377
Db 4038 TGTGATTTGGGGCTGATGGGAAACCTCCCTGGAGCAAGTGTGCTCCCTGCAATGCTCC 4097
Qy 1377 oCysGlyGlyGlnTyrValGlySerAspGlyValValLeuSerProAsnTyrProGlnAs 1397
Db 4098 CTGTGAGAGCCAGTATCACCGGATCAGAAAGGGATGTTTATGACCAAACTACCCCATGA 4157
Qy 1397 nTyrThrSerGlyGlnIleCysLeuThrPheValThrValProLyAspTyrValValPh 1417
Db 4158 TTTACAGCTGTGCAAAATATGCTCTATTCATCACGGTACCAAGAAATGTGTGCTTT 4217
Qy 1417 eGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspValValGluValHisAspGlyHis 1437
Db 4218 TGGACAGTTTGTCTATTTCCAGACAGCCCTGATATTTGGCAGAAATTTGATGAGAAC 4277
Qy 1437 sSerGlnHisSerAlaGlyLeuLeuSerSerLeuSerGlySerHisThrGlyGluSerLeuP 1457
Db 4278 CCATGCACAGGCCAGACTTCTCAGCTCACTCCGGGTCTCACTCAGGGGAAACATTCGC 4337
Qy 1457 oLeuAlaThrSerAsnGlnValLeuIleLysPheSerAlaLysGlyLeuAlaProAlaP 1477
Db 4338 CTTGGTACGTCAATCAATTCATCTGCTCCGATTCAGTGCAAAGACGGTCTTGCCCG 4397
Qy 1477 gGlyPheHisPheValTyrGlnAlaValProArgThrSerAlaThrGlnCysSerSerVa 1497
Db 4398 CGGCTTCCACTTCGTGATCAAGCTGTCTGTGACAGTACCAACATGACGCTGT 4457
Qy 1497 lProGluProArgTyrGlyLysArgLeuGlySerAspPheSerValGlyAlaIleValAla 1517
Db 4458 CCCCAGGCCAGATACGAAAGAGAAATGTTGTGAGTTTCTGCGGCTCATGTGCTCG 4517
Qy 1517 gPheGluCysAsnSerGlyTyrAlaLeuGlnGlySerProGluIleGluCysLeuProVa 1537
Db 4518 ATTCAGAGGCAACCGGGATACCTGCTTCAGAGTTCCACGGGCTCCACTGCAATCCGT 4577
Qy 1537 lProGlyAlaLeuAlaGlnTrpAsnValSerAlaProThrCysValValProCysGlyG 1557
Db 4578 GCCCAACCCCTTGGACAGTGGAAACGACACGATCCCAAGCTGTGTGTAACCTGAGTGG 4637
Qy 1557 yAsnLeuThrGluArgArgGlyThrIleLeuSerProGlyLysPheProGluProTyrLeuAs 1577
Db 4638 CAATTTCACTCAACGAAAGATCAATCTGTGCCCCCGGCTACCTGTAGCCATACGAGAA 4697
Qy 1577 nSerLeuAsnCysValTrpLysIleValValProGluGlyAlaGlyIleGlnIleGlnVa 1597
Db 4698 CAATTTGAACGTATATGAAAGATCATATGTTACGAGAGGCTCGGGAATTCAGATCCAGT 4757
Qy 1597 lValSerPheValThrGluGlnAsnTrpAspSerLeuGluValPheAspGlyAlaAspAs 1617
Db 4758 GATCAGTTTTCACAGGAGCAAGAACTGGGACTCCCTTGAGATTCACAGATGTGTGGGATGT 4817
Qy 1617 nThrValThrMetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerTh 1637
Db 4818 GACCGCACCCAGACTGGAGACCTTCTCAGGACACAGATACCGGACCTGCTGAACAGATAC 4877
Qy 1637 fSerAsnGlnLeuTyrLeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHis 1657
Db 4878 TTCCAAACCACTACCTGACGATTTCCAGTGCATTAATGATGACACTGCTGATTTCCA 4937
Qy 1657 eLeuGluTyrTrpThrValGlyLeuSerSerCysProGluProAlaValProSerAsnGl 1677

```

Db 4938 CCTGAATACAAAAGTGTAGTCTTGTCATGCCAAGACCGCTCCCGAACAAG 4997
 Qy 1677 yVallyserThrGlyGluValThrLeuValAsnAspValValSerPheGlnCysGluProG1 1697
 Db 4998 CATCAAAATCGAGATCGGTACATGTGTACAGACGATGCTCTCCAGTCGAGCCGG 5057
 Qy 1697 yTAlaLeuGlnGlyHisAlaHisIleSerCysMetProGlyThrValAlaGAGTTPAs 1717
 Db 5058 GTACACCTTCAGAGGCGCTTCCACATTTCTGTATGCGACGAGCCGCTCCGTTGAA 5117
 Qy 1717 nTyProProProLeuGlyHisAlaGlnCysGlyGlyThrValGluGluMetGluGlyVa 1737
 Db 5118 CTATCCGTCCTCCCTGTGATTCGAACTGTGAGAGGAGCGCTGAGCACTTGGGGTGGT 5177
 Qy 1737 lIleLeuSerProGlyPheProGlyValnTyProSerMetLeuPheCysSerTPylsI1 1757
 Db 5178 GATCCGTGAGCCCGGCTCCAGGTTCTTACCCCAACACTTACCTGACCTGGAGAT 5237
 Qy 1757 eAlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheserThrGluProAs 1777
 Db 5238 CTCATTCACCATTCGGCTATGTGTGACATATTCAGTTTCTGAATTTTCTACCGAAGCTAA 5297
 Qy 1777 nHisAspTyxIleGluIleArgAsnGlyProTyGluThrSerArgMetMetGlyArgPh 1797
 Db 5298 TCATGACTTCTCTGAATTCAAAATGAGACCTTACACACACGAGCCCATGATGGACATTT 5357
 Qy 1797 eSerGlySerGluLeuProSerSerLeuLeuSerThSerHisGluThrValTyxPh 1817
 Db 5358 TAGCGGACGAGATCTCCCGGCGCTCTGAGCACAACGATGAAACCCATCATCCACTT 5417
 Qy 1817 eHisSerAspHisSerGlnAsnArgProGlyPheLeuLeuGluTyxGlnAlaTyxGluLe 1837
 Db 5418 TTATATGATACATTCGCAAAACCGCAAGATTTAACTTGCTTACCAAGCCCTATGAAT 5477
 Qy 1837 uGlnGluCysProAspProGluProPheAlaAsnGlyIleValArgGlyAlaGlyTyxAs 1857
 Db 5478 ACAGAACTGTCCAGATCCACCCCATTTCCAGATGGATGATCAATCACTCGGATTAAG 5537
 Qy 1857 nValGlyGlnSerValThrPheGluCysLeuProGlyTyxGlnLeuThrGlyHisProVa 1877
 Db 5538 CGTGGGGCAATCAGATATCTTTCGAGTGTATCTGGGTATCAATTCAAAGGCCATCTCGT 5597
 Qy 1877 lLeuThrCysGlnHisGlyThrAsnArgAsnThrAspHisProLeuProLysCysGluVa 1897
 Db 5598 CTTCACTTGTCAAGCATGGGATCAACAGAACTGGAACTAACCTTTTCCAAAGATGGAGCC 5657
 Qy 1897 lProCysGlyGlyAsnIleThrSerSerAsnGlyThrValTyxSerProGlyPheProSe 1917
 Db 5658 CCTTGTGGGTACAAAGTAACTTCTCAGAACGGCACATCTACTCCCTGGCTTCCGTA 5717
 Qy 1917 rProTyxSerSerSerGlnAspCysValTyxPheLeuIleThrValProIleGlyHisGlyVa 1937
 Db 5718 TGAGATATCCGATCTCGAAGGAGCTGATTTGGCTCATCAAGGTGCTCCAGGGCACGGAGT 5777
 Qy 1937 lArgLeuAsnLeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTPAs 1957
 Db 5778 TTACATCAACTTCACCTGTATACAGACGGAAGCTGTCAACATTTACATTCCTGTTGGGA 5837
 Qy 1957 pGlyProGlnGlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaTyxLeuTh 1977
 Db 5838 CGGTGCCGATCAAGAACTCAACCCGAGCTGGAGTTTTCAGTGGCAACACAGCCCTCGAAG 5897
 Qy 1977 rValGlnSerSerSerAsnGlnValLeuLeuLeuPheHisIleArgAspAlaIleThrGlyG1 1997
 Db 5898 GGGGTATAGCTCCCAACCAAGTCTGCTCAAGTTCCACAGCGCACTTTCAAAATGAGAG 5957
 Qy 1997 yIlePheAlaIleAlaPheSerAlaTyx-----ProLeuThrLys----- 2010
 Db 5958 CTTCTTGTCTCAATTTCCACGGTCACTGATTTTCACTCCGTTAGTTAAAGACTGAGAA 6017
 Qy 2011 -----CysProProPro-----ThrIleLeuPr 2018
 Db 6018 TTCCATGTGTGTTTACTGACAGTGTGTCTCCACGCTTGTTCACGATGAAGTTTCTTGA 6077

Qy 2018 oAsnAlaGluValValThrGluAsn 2026
 Db 6078 TTCAGCCGAGGCGGTATGATTTCT 6102
 RESULT 7
 US-10-276-934-1
 ; Sequence 1, Application US/10276934
 ; Publication No. US20030180750A1
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Leeds
 ; APPLICANT: Markham, Alexander F.
 ; APPLICANT: Jackson, Andrew P.
 ; APPLICANT: Woods, Christopher G.
 ; TITLE OF INVENTION: Treatment of Cancer and Neurological Diseases
 ; FILE REFERENCE: 9052-144
 ; CURRENT APPLICATION NUMBER: US/10/276,934
 ; PRIOR FILING DATE: 2002-11-20
 ; PRIOR APPLICATION NUMBER: PCT/GB01/02240
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: GB0012186.3
 ; PRIOR FILING DATE: 2000-05-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 5598
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-276-934-1
 Alignment Scores:
 Pred. No.: 0 Length: 5598
 Score: 7159.00 Matches: 1262
 Percent Similarity: 81.17% Conservative: 264
 Best Local Similarity: 67.13% Mismatches: 279
 Query Match: 42.15% Indels: 75
 DB: Gaps: 7
 US-10-016-248-2 (1-3104) x US-10-276-934-1 (1-5598)
 Qy 36 VallyserGlnIleGluLeuLysSerArgGlyVallyserMetProSerLysAspAsn 55
 Db 66 GTGAAAAAGCGATGAGTTGATGATCAAGAGAGTCAAGAGTGTGCCAGCAAGATGGA 125
 Qy 56 SerGlnTyxThrSerVal----- 61
 Db 126 AGCCATAAAACTGTGCTGTGCAATGATGAAATACATCAACCATGTGATCGAAGAAGA 185
 Qy 62 -----LeuThrGln-ValGlyValSerGlnG1 70
 Db 186 TACGTTATTATGTATCTGGGAAAAACCATGGCTTCAAGTGAATTCAGTGAAGCCAAAG 245
 Qy 70 y-----HisAsnMetCysProAspProGlyIleProGluArgGlyTyxAs 85
 Db 246 AGGTGTTCATGATGTCTGTGACATGTGTCCAGATCTCGGATTCAGAAAAATGTGAGAG 305
 Qy 85 gLeuGlySerAspPheArgLeuGlySerSerValGlnPheThrCysAsnGluGlyTyxAs 105
 Db 306 AGCAGGTTCCCACTTCAGGCTGGTGTGCAAAAGTACAGTTTTCATGTGAGCAATTAAGT 365
 Qy 105 pLeuGlnGlySerTyxArgIleThrCysMetLysValSerAspMetPheAlaIleTPSe 125
 Db 366 GCTCCAGGATCTTAAAGCATCTGCTGACAGAGATTACAGAGAGCTCGCTGCTGGAG 425
 Qy 125 rAspHisArgProValCysArgAlaArgMetCysAspAlaHisIleLeuArgGlyProSerG1 145
 Db 426 TGACACAGCGCCCATCTGCGAGGAGGAACATGTGATCCAAATCTGGGTGGCCCAAGCG 485
 Qy 145 yIleIleThrSerProAsnPheProIleGlnTyxAspAsnAlaHisIleCysValTyxP1 165
 Db 486 CGTCATTAACCTCCCTTAATTAATCCGTTTCACTATGAAGATTAATCACACTGTGTGGGT 545
 Qy 165 eIleThrAlaLeuAsnProSerLysValIleLysLeuAlaPheGluGluPheAspLeuG1 185

```

Db      546 CATCACCAACACCCGACCAAGTCATCAAGCTTCCTTTGAAGGTTTGAGCTGGA 605
Qy      185 UARGLYLTYRAPHTRLEUTHVALGLYASPGLYGLYLSAPGLYASPGLYLSTHVA 205
Db      606 GCGAGGCTATGACACCTGACCGTGTGTGATGCTGGGAAGGTGGAGACACCAATCGGT 665
Qy      205 LLEUTYRMESESGINAEALACYSERASPSERPROHISTHPRGGLYSEARAIIIEPR 225
Db      666 CTTCGACGTG-----CTCACGGGATCCAGTGTTC 695
Qy      225 OGLU-----SERMETSERGLYASPILETTPARGLINLSTPTHTVALLEUGLIU1 242
Db      696 TGACCTCATGTGAGCATGAGCAACCAAGATGCGTACAT----- 735
Qy      242 ECYARHASPILLESERSESERASPAALARGSERGLYSERVALARGLYSERPROLYSTH 262
Db      736 -----CTGCAGTGTGATGATGATGAC----- 753
Qy      262 RSEASNAIAYALGLUEUVALAIPROGLY-----THRGUILEGLUG1 277
Db      754 -----ATTGGCTCACTGGGTTTAAAGCTGTTACCAAGAAATTGAAA 797
Qy      277 NGLYSERCYSGLYASPPROGLYIILEPROALATYRGLYARGARGGLUGLYSERARGPHEH1 297
Db      798 GGGAGGGGTGTGGGATCCTGGAAATCCCGCTATGGGAAGCGGACGGGACAGATTTCCT 857
Qy      297 SHAGLYASPTHTLEUYSRPHGGLUCYSGINPROALAIPHEGLIUEUVALGLYGLINLSA1 317
Db      858 CCAITGGAGATACACTCACCTTTGAAATGCCGCGGCTTTGAGCTGTGTGGGAGAGAGT 917
Qy      317 ALIETHCYSGINLUSASNAENGINTPSEARALALYELASPROGLYCYSAVALPHESECY 337
Db      918 TATCACCCTGTGACAGAACATCATGCTGTGTGGCAACAGCCAGCTGTGATTTTCATG 977
Qy      337 SPHEPHEASNPHEHTRHSEPRROSERGLYVALVALLEUSERPROASNTYRPROGLUASPTLY 357
Db      978 TTTTTCACCTTAACTTACCGCATCATCTGGGATTTATCTGTCAACCAATTATCAGAGGAAATA 1037
Qy      357 RGLYASNHILEUHSICYSVALTTPLEUILEUVALARGPROGLUSERARGILEHISTE 377
Db      1038 TGGGAACAACATGACTGTCTGTGTTGATTTATTCGAGCCAGGAAGTGAATTCACCT 1097
Qy      377 UALAPHEASNAAPILLEASPYALGLUPROGINPHEASPPHELEUVALILEYASPGLYA1 397
Db      1098 AATCTTAATGATTTGATGTGAGCTCAATTGACTTTCTCGCGGTCAAGGTAGTAGTG 1157
Qy      397 ATHRALGLUALAPROVALLEUGLYTHRPHSESGLYASNGINLEUPROSERILETH 417
Db      1158 CATTTCTGACATTAATCTGTCTGGGTACTTTTTCGGCAATGAAGTGCCTTCCACGCTGGC 1217
Qy      417 RSESESGLYHIEVALALARGLEUGLUPHEGINTHRASPHISESTHGLIYLSARGI 437
Db      1218 CAGAGAGGGGCAATTAAGTTCGTTGGAATTCAGTCTGACCACTTCCACTCTGCGCAGAG 1277
Qy      437 YPHEASNHILETHPHEHTRHTRPHEARGHISANGIUCYSPROASPPROGLYVALPROVA 457
Db      1278 GTTCACATCATCTTACACCACTTTGCTCAGAAATGAGTGCATGATCTCGCATTCCTAT 1337
Qy      457 LAANGILYASARGPHEGLYASPSERLEUGINLEUGLYSERSETHLESETHLEUYSAS 477
Db      1338 AAACGAGACGAGTTTGGTGAACAGTTTCTACTCGGAGGCTGGGTTCTTTCACATGTGA 1397
Qy      477 PGLUGLYPHELEUGLYTHRGINLYSERGLUTHRIETHCYSAVALLEUYSGLIYSE 497
Db      1398 TGAATGGCTTGTCTGAAGCCAGGGATCCGATCATTAAGTCTGCAATCTGCAACAGGGA 1457
Qy      497 RVAVALITTPASERVALVALLEUARGCYSGIUALAPROCYSGIYGLYHISLEUTHISE 517
Db      1458 CGTGGTGTGAGCTCCACCGTGGCCGCTGTGAAGTCCCATGTGGTGACATCTGCACAG 1517
Qy      517 IPRROSERGLYTHRIELEUSERPROGLYTPPROGLYPHETRYLYASPAALALEUSERCY 537
:      :
Db      1518 GTCCAGGAGATCATTTTGGCTCTCGATGCGCCAGATATTATAAGATTCTTTACATTG 1577
Qy      537 SALATRPVALILEGLUALAGINPROGLYTYRPROILEUSILETHPHEASPARPHEU 557
Db      1578 TGAATGATATTAAGAACAAACCAAGCCACTCTATCAAAATATTAATCTTTGACAGATTTC 1637
Qy      557 ETHRGILVALASNTYRAPHTRHLEUGLUALARGASPGLYARGTHTRYSERLAPROLE 577
Db      1638 GACAGAGGTCATTAATTAAGACCTTGGAGGTCAAGATGGGCCACGACGATTCGCCACACT 1697
Qy      577 ULILEGLYVALTYRHISGLYTHRGINLVALPROGINPHELEUILESERTHSERANTYRLE 597
Db      1698 GATGGCGAGTACACAGGACCCAGGACCCAGATTCTCATAGACCCGGGAATTCAT 1757
Qy      597 UTYRLEULEUPHESETHASPLYSESTHISERASPIILEGLYPHEGLINLEUARGTYRGI 617
Db      1758 GTACCTGCTAATTCACACATGACACCAAGCCGCTCCAGCATCGGCTTCCTCATCATATGA 1817
Qy      617 UTHRIETHTRLEUGINSESRASPHISECYSELEUASPPROGLYIILEPROVALASNGIYGLINR 637
Db      1818 GAGTGTGACGCTTGAAGTGGATTCCTGCTGGAACCGGCGATCCCTGTAAAGGCGCATCG 1877
Qy      637 GHISGLYASNAAPPHETRYVALGLYALALEUVALTHRPHSERCYASPSERGLYTYRTH 657
Db      1878 CCAAGGTGAGACTTTGGCATCAGGTCCACAGTACTTTCACTGTGACCCGGGGTATCAC 1937
Qy      657 ILEUSERASPGLYGLUPROLEUGLUCYSGIUEUROASNPHEGINTPSERARGALALEUDR 677
Db      1938 ACTAAATGACGACGAGCCCTCGTGTGAGAGGAACCAACGATGAMACACGCTTGGCC 1997
Qy      677 OSERCYSGIUALALEUCYSGIYGLYPHEILEGLINLYSESERGLYTHRIELEUSERP 697
Db      1998 CAGCTCCAGCGCTCATGTGGAGGCTCATCTCAAGGAAGATGGAACAGTCTTTCTCC 2057
Qy      697 OGLYPHEPROASPPHETYRPROASNALEUASNCYETHTRIPRIELIEGLIUTHSEXH1 717
Db      2058 TGGGTTTCCAGATTTTATTCACAACTCTCTAATGACAGTGGACCATTTGAAGTGTCTCA 2117
Qy      717 EGLIYASGLYVALPHEPHEHTRPHEHISTHTRPHEHISTLEUGIUSERGLYHISAPTYRLE 737
Db      2118 TGGGAAGAGATTCAATGATCTTTCACACCTTTATCTGTGAGATTCCCAAGACTATTT 2177
Qy      737 ULEUILETHRGILUANGIYSESRPHEHTRINPROLEUARGINLEUTHRGILYSEARGLE 757
Db      2178 ACTGATCAACAGAGATGAAATTTTCCGAGCCCTTTCGACGCTCACCGGTCGGTGT 2237
Qy      757 UPROALAPROILESERALAGLYLEUITYRGLYASNPHEHTRALAGINVALARGPHEILESE 777
Db      2238 GCTCATACGATCAAGGACGAGGCTGTGTGAAACTTCACTGCCAGCTTCGGTTATATTC 2297
Qy      777 RASPPHESEMETSETYRGLUGLYPHEASNHILETHRPHSESGIUTYRASPLEUGLUPR 797
Db      2298 AACCTTCTCAATTTGTGAGAGGCTTCAATATCAATTTTCAGAAATGACCTTGAGGC 2357
Qy      797 OCYSGIUGLUPROGLIUALPROALATYRSETHLEARGLYSGIYLEUGINPHEGLYVALGI 817
Db      2358 ARGATGATATCTGTGAGTCCGCTTCAGCGCAAGAAATGTTGTTTCACTTTGAGTGGG 2417
Qy      817 YASPTHTLEUTHTRPHESECYSPHEPPOGLYTYRARGLEUGLUGLYTHRIALARGILETH 837
Db      2418 AACCTCTGACGTTTCTCTCTCTGAGATTCGTGTAAAGGTGCACCAAGCTTAC 2477
Qy      837 CYSELEUGIYGLYARGARGLEUTRPSERSEPRROLEUPROARGCYSAVALAGIUCY 857
Db      2478 CTGCTGTGGTGGGGCCCGCTGTGTGAGATGACACTTTCGCAAGGTGTGGCCGAATG 2537
Qy      857 EGLYASNSERVALTHRGIYTHRGINGLYTHRIELEUSERPROASNPHEPROVALASNTY 877
Db      2538 TGGAGCAAGTGTCAAAAGAAATGAAGAAATTAATCTGCTCCAAATTTTCCATCAATTA 2597
Qy      877 RASASNAHNSHISGLUCYSILETYRSETHLEGINTHRGINPROGLYLSGLIYGLINLE 897
Db      2598 TGAATATAACATGATGATATTAATAATGAAGAACGCGCAAGGGCATTCACACT 2657

```

QY 897 ULVLAARGLAAPHGIEULEUSERGLUGLYASPVALLEUUYVALTYRASPGLYAAMS 917
DB 2658 TAGAACACGAGAGCTTCACAGCTGTTTGAAGAGATACCTAAAGATATATGATGAGAAAAGA 2717
QY 917 NAENSERALAARGLEULEUGLYVALPHESETHISERGJLUMETWEGLYVALTHRLAUS 937
DB 2718 CAGTTCCGTCACGTCACAGTGGCAGGTTACATAAAATGAACTTCGGGGCTGATCTTAA 2777
QY 937 NSETHISESERSEULEUTPLEUAPHPHEILETHASPLAAGLUANTHRSERLYSGI 957
DB 2778 CAGACACATCCCAATCACCTGTGCTAGATTCAACACCAATGATCTGCACCGACCAAGG 2837
QY 957 YPHGIEULEUHIAPHESESERPHEGIEULEUILEUYSGEGLUAPPROGJLYTHRPROLY 977
DB 2838 TTTTCACCTCACCTATACAGTTTATGATCTGTATAAATGATGAGATCCGGGCATCTTAA 2897
QY 977 APHEGLTYRLYVALHISAEPGLUGLYHISAEPALAGLYSESERVALSERPHESECY 997
DB 2898 CTACGGGCTATAGATCCGCTGATGAAGGCCACTTTACGACACATGATCTGTACAGTTG 2957
QY 997 SAEPFPROGJLYTYRISERLEUARGLYSERGLUGLEULEUCYSEUSERGLYGLUARGAR 1017
DB 2958 CAACCCGGGCTAACCCCATGCGACATGCGACACACCCCTGACCTGTTGAGTGAGACAGAG 3017
QY 1017 GTHRTTPASPARGPROLEUPROTHRCYVALAAGLUCYSGEGLYGLYTHRVALLAARGLYGI 1037
DB 3018 AGTGTGGACAAACCACTACCTTCGTGATGCGGAATGTGTGTGATGATGATGATGATGATG 3077
QY 1037 UVALSERGLYGLINLLEUSERPROGLYTYRPROALAPROTYRGLUHIASAENLEUAENCY 1057
DB 3078 CACATTCAGAGACGATATGTCCTCGCTATCCAGCTCCGATATGACAAACACCTCCACTG 3137
QY 1057 SILETTPHRTILEGLUALAGLYVALAGLYCYSTHRTILEGLYUHIAPHELEUVALPHEAS 1077
DB 3138 CACCTGGATTATPAGAGGACAGACCAAGAAACCATTTAGCTTCATTTGATTTTGA 3197
QY 1077 PTHRTGLUGLUVAHISAEPVALLEUARGILETPASPGLYPROVALGUSERGLYVALLE 1097
DB 3198 CACGAGAGTGGCTCACGACATCTTCAGTGTGGAGCGGGCCGGGACATGACATCTCT 3257
QY 1097 ULEULYGLIULEUSERGLYPROALALEUPROLYASPLEUHIASERTHPHEANSEVA 1117
DB 3258 GCTAAGAGAGTGGCTGCTCCGCCCTTCGGAGAGACATCCACAGCACCTTCACTACT 3317
QY 1117 LVALLEUGLNPHSETHRASPHEPHERTHSERISGLINGLYPHEALALEGINPHESE 1137
DB 3318 CACCTGCGATTGACAGACGACTTCTTCATCAGCAAGTGTGCTTCATCCAGTTCTC 3377
QY 1137 FVALSERTHRALAHRSECYSAENAPPROGLYILEPROGLINENGLYSEARAAGSERGI 1157
DB 3378 CACCTCAATTGCAACCCCTGTAAAGATCCAGTATGCCCAAAATGGCACCCCTATGG 3437
QY 1157 YASPSETTRPGLUAGLYASPSETTRVALPHEGLINCYASAPPROGLYTYRVALALEUGI 1177
DB 3438 AGAAGACGAGAGGCTGAGACACCCGTACATTCAGTGTACCTGCTATCAGCTCA 3497
QY 1177 NGLYSERALAGLULESECYSAVALYSEILEGLUASNAHRPHEPHERTPGINPROSEPR 1197
DB 3498 AGGACAAGCCAAATACCTGCTGTGACAGCTGAATPACCGGTTCTTTGGCAACGACACC 3557
QY 1197 OPRTHCYVALLEALAPROCYSGEGLYASPLEUTHRCYPROSEGLYVALILELEUSE 1217
DB 3558 TCTTCAATGCAATGCTCTTGTGAGGGAATCTGACCGGGCCACACAGGTGTATTTTGTCTC 3617
QY 1217 RPROASNTYRPROGLUPROTYRPROPROGLYLYSGIUCYASAPTRPLYSVALTHRVASE 1237
DB 3618 ACCCACTAACCCACAGCCGTATCTCTCTGGAGAAGAAATGTACGTGAGAGATGAAGTCAA 3677
QY 1237 RPROASPTYRVALILEALALEUVALPHEASNILEPHEANLEUGLUPROGLYTYRASPHE 1257
DB 3678 CCCGACATTGTGTCTGCTGTGATATCAAAAGTTTCAACATGAGCCCACTATGACTT 3737

QY 1257 ELEUHIILETYRASPGLYARGASPSETLEUSERPROLEUILEGLYSERPHETYRGLYSE 1277
DB 3738 CTTACACATCTATAGAGGAGAAAGTTCCAAACGCCCTCTATTGGAGTTTACAGGCTC 3797
QY 1277 RGLINLEUPROGLYARGILEGUSERSEASNSERLEUAPHELEUALAPHEARGSERAS 1297
DB 3798 TCAGGCCCCGAAAGATPAGAGTACGCGAAAACAGCTGTTTCTGCGCATTTCCGAGTGA 3857
QY 1297 PALASERVALSERNAAGLYPHEVALILEASPTYRTHGLUASNPROARGJLUSERCY 1317
DB 3858 TGCTCTCGTGGGCTTTCAGGGGTTCGCCATTGAATTTAAAGAAACCAACCGGAGAGCTTG 3917
QY 1317 SPHEASPPROGLYSERTILEUYSAENGLYTHRARGVALIGLYSEASPLEUUYSLAUGLYSE 1337
DB 3918 TTTTGACCCGAGAAATATATGATGAGACAGAGTTGGAAACAGACATTTCAAGCTTGCTC 3977
QY 1337 RSEVALTHRTYRRCYSHISGLYGLYRGLUVALIGLUGLYTHRSETHRLSEUSERCY 1357
DB 3978 CACCATCACCTACAGATGATGATCTGCTGCTATTAAGATTCTTGACCCCTCATCATCAGCTG 4037
QY 1357 SILELEUGLYPROASPGLYLYSPROVALTPPASNPAPARGPROVALCYSTHRTIAPR 1377
DB 4038 TGTGATTTGGGCTGATGGGAAACCCCTCGGACCAAGTGTGCTCTCTGCAATGTCTCC 4097
QY 1377 OCYSGLYGLYGLINTYRVALIGLYSERASPGLYVALILEUSERPROASNTYRPROGLINAS 1397
DB 4098 CTGTGAGGCGCAGTACACGGGATCAGAAAGGGGTGTTTATCACCAACTACCCCATBA 4157
QY 1397 NTYRTHSERGLYGLINLLECYSELYTYRPHIVALPHEVALPVALPROLYASAPTYRVALIAPH 1417
DB 4158 TTACACAGCGTGTCAAAATATGCTCTATTCATCAGCGTATCAAAAGAAATTCGTGCTT 4217
QY 1417 EGLYGLNPHEALAPHEPHERTHRTALALEUASAPVALAIGLUVAHISAEPGLYHI 1437
DB 4218 TGACAGTTTGGCTCATTTTCAGACAGCCCTGAAGATTTGGCAAAATATTATGAGAAC 4277
QY 1437 SSERGLNHIASERARGLEUSERSESERLEUSERGLYSETHIRHGLYGLUSERLEUPR 1457
DB 4278 CCATGCACAGGCCAGACTTCTCAGCTCACTCTCGGGGTCTCATCAGGGGAAACATTGCC 4337
QY 1457 OLEUALATHSERANGLIVALLEUULEUYSPHESERALIYSGLYLEUALAPROALAPR 1477
DB 4338 CTGTGCTACGTCAAATCAAAATTCGTCTCCATTCAGTGAAGAAAGCGGTGCTCGCCCG 4397
QY 1477 GGLYPHEHIAEPHEVALYRGLINALAVALPROARGTHSERALATHRGLINCYSESERVA 1497
DB 4398 CGGCTTCACATTCGTATCAAGCTGTTCCTCGTACAGTGAACCAATGCACCTCTGT 4457
QY 1497 LPROGLUPROARGTYRGLYASARGLEUGLYSEASPPHESEVALIGLYALILEVALAR 1517
DB 4458 CCCCAGCCCGACGATACGAGAGAGAAATGGTTCTGAGTTTCTCCGGCTTCATGTCCG 4517
QY 1517 GPHGIEUCYASNSERGLYTYRVALALEUGLNGLYSEPROGLIULEGLUCYSELEUPROVA 1537
DB 4518 ATTGAGATGCAACCCGGGATATCCTGCTTCAAGGTTTCCACGCGCTCCAGCTCAGTCCGT 4577
QY 1537 LPROGLYALALEUALAGLINTPASPENVALSERALAPROTHRCYSAVALYALPROCYSGLYGI 1557
DB 4578 GCCCAAGCGCTTGGACAGTGGAAACGACACATCCAGAGCTGTGTGTAACCTCCAGTGG 4637
QY 1557 YASNLEUTHRGUARGARGLYTHRTILEUSERPROGLYPHEPFGIUPROTYRLEUAS 1577
DB 4638 CAATTTCACCTCAACGAAAGATGATCAATCTGTCCCGGGCTACCTGAGCCATACGGA 4697
QY 1577 NSERTLEUASNYVALTRPLYSVALILEVALIAPROGLUGLYALAGLYILEGINLLEGINVA 1597
DB 4698 CAACTTGAATGATATGAGAAAGATCATAGTTACGAGAGGGCTCGGGAATTCAGATCAAGT 4757
QY 1597 LVALSERPHEVALTHRGLUGLINTPASPSETLEUGLYVALPHEASPGLYALAPAPAS 1617
DB 4758 GATCAGTTTGGCCAGGAGAGAACTGGACTCCCTTGAATTCACACATGTGTGGGAGTGT 4817
QY 1617 NTHRTVALTHMETLEUGLYSERPHESEGLYTHRTVALPROALALEUASNSERTH 1637

```

Db      4818 GACCCACCCAGACTGGAGCTTCTCAGGACCACTACCGGACTGCTGAAACAGTAC
      1637 rSerasnGlnLeuTyrLeuHisPheTyrSerAspIleSerValSerAlaIaIaGlyPheH
      4878 TTCACCAACCTTACCTGACATTTCTCAGCTGACATTTGAGTGTGGAGCTGCTGTTTCCA
      1657 sLeuGlnTyrIleThrValGlyLeuSerSerCysProGluProAlaValProSerAsnG
      4938 CTTGGAAATCAAAATCTGATAGCTTCTGCTGACATGCCAAGAACAGCCCTCCACCAACAG
      1677 yValIleThrGlyGluATGTYrLeuValAsnAspValIleSerPheGlnCysGluProG
      4998 CATCAAAATCGAGATCCGTCATGTCATGTAACGACGTGCTCTCTCCAGTCCGAGCCGG
      1697 YTrAlaLeuGlnGlnGlyHisAlaHisIleSerCysMetProGlyThrValAlaGAGTTPAs
      5058 GTACACCTGACGGGCGCTTCCACATTTCTGATGCGACGAGACCGCTCCGCTTGGAA
      1717 nTyrProProProLeuGlyCysIleAlaGlnCysGlyGlyThrValGluGluMetGluGlyVa
      5118 CTATCCGCTCTCCCTGCTGATTCGACAACTGTGAGAGGACGCTGACACCTTGGGTGTGT
      1737 lIleLeuSerProGlyIlePheProGlyAsnTyrProSerAsnMetAspCysSerTyrIle
      5178 GATCCTGAGCCCGGCTTCCAGGCTTCTTACCCCAACAACTTACGACTGACACTGAGAGAT
      1757 eAlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGluProAs
      5238 CTCATTACCATCCGCTGATGTCATGTCATATTCATGATTTCTGATTTCTACCGAAGCTAA
      1777 nHisAspTyrIleGluIleArgAsnGlyProTyrGluThrSerArgMetGlyArgPh
      5298 TCATGACTTCTCTGAAATTCAAATGACCTTACACACCGAGCCCATGATGAGCATTT
      1797 eSerGlySerGlnLeuProSerSerIleuSerIleuSerThrSerHisGluThrThrValTyrPh
      5358 TAGGGGACGAGATCTCCCGGGGCGCTGTCGACACACGATGAAACCTCTCATCCCACTT
      1817 eHisSerAspHisSerGlnAsnArgProGlyPheTyrLeuGluTyrGlnAlaTyrGluLe
      5418 TTATAGTACCATTCGCAAAACCGGCAAGATTTAACTTCTTACCAAGCTTATGATTA
      1837 uGlnGluCysProAspProGluProPheAlaAsnGlyIleValArgIleValIaGlyTyrAs
      5478 ACAGAACTGTCCAGATCCACCCCATTTCAAGATGGTACATGATCAACTCGGATTACAG
      1857 nValGlnGlnSerValThrPheGluCysLeuProGlyTyrGlnLeuThrGlyHisPro
      5538 CGTGGGGCAATCAATATCTTTCAGTGTATCTCGGGTACATTCATATAGGCATCCT
      5595

```

```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (588)..(588)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (624)..(624)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (672)..(672)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (2265)..(2265)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (5468)..(5468)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc feature
LOCATION: (6027)..(6027)
OTHER INFORMATION: "n" is any nucleotide
US-10-276-934-2
Alignment Scores:
Pred. No.      0
Score:         6991.00
Percent Similarity: 80.90%
Best Local Similarity: 66.83%
Query Match:    41.16%
DB:             15
Gaps:           7
US-10-016-248-2 (1-3104) x US-10-276-934-2 (1-6145)
      36 ValIleValGlnIleGluLeuTyrSerArgIleValIleLeuMetProSerIleAspAsn
      66 GTGAAATAAGCGCATTTGATGTAAGTCAAGAGATCAAGATCTGCTCCGCAAGATGGA
      56 SerGlnIleThrSerVal
      126 AGCATTAATAACTCTGTCTGGCATATGAAATACATCAACATGTGATCCGAAAGAGA
      62 -----LeuThrGln-ValGlyValSerGlnI
      186 TAGCTTCAATATATCTGGGAAACCATGCGCTTCAAGTGAATTCAGTACGACCAAG
      70 Y-----HisAsnMetCysProAspProGlyIleProGluArgIleValArg
      246 AGGTGTGATGCTGCTCTGACATGTCTCCAGATCTCGGATTTCCAGAAATGTATGAAG
      85 gLeuGlySerAspPheArgLeuGlySerSerValGlnPheThrCysAsnGluGlyTyrAs
      306 ACCAGTTCGCACTTCAGGTTGTGCAAAATGTACATTTCACTGATGAGCAATTAAGT
      109 pLeuGlnGlySerIleValGlyIleThrCysMetIleValSerAspMetPheAlaIaIaTyrSe
      366 GCTTCAGGAGATCTAAACATCACTGTCAGAGATTACAGAGAGCTTCGCTGCTTGAGG
      125 rAspHisArgProValCysArgAlaArgMetCysAspAlaHisIleuArgGlyProSerG
      426 TGACACAGGCCCATCTCCAGACGAAACATGTGATCCAACTGCTGCGGCCACAGCGG
      145 yIleIleThrSerProAsnPheProIleGlnTyrAspAsnAlaHisCysValIleTyrI
      486 CTTCAATTCCTCCCTAATTAATTCGGTTCAAGATGATATGACACCTGTGTGGGT
      165 eIleThrAlaLeuAsnProSerIleValIleIleValLeuAlaPheGluGluPheAspLeuG
      546 CATCACCAACCAACCGGACAAAGTATCATCAAGCTTCCCTTNGAAGATTGACCTGA
      605

```

QY	185	IARGGLVYTRRPTTRTTRTTRVVALGLYASRGLYGLYGLNMRGLYASRGLNLYSRTHVA	205
DB	606	GGAGGCTATAGCACCTTACGGTGGTGAATCGTGGAGAGTGGAGACACACAGATCGGT	665
QY	205	lLeuTYmetSerGlnAmaAlaCySerAspSerProHlSRProGlySerArgLlePr	225
DB	666	CTTGATGATG-----CTCAGGGATTCAGGTTC	695
QY	225	oGlu-----SerMetSerGlyAspLleTPArgGlnLYSRTHrVAlleuGluI	242
DB	696	TGACCTCATTTGACATGAGCAACCAAGATGCTACAT-----	735
QY	242	eCyArgAspLleSerSerSerAspAlaArgSerGlySerValArgLYSerProLYSRTh	262
DB	736	-----CTGAGCTCGATGAATGAC-----	753
QY	262	rSerAmaAlaValGluLeuValAlaProGly-----ThrGluIleGluI	277
DB	754	-----ATTGGCTCACCTGGGTTAAACTGTTTACCAAGAAATTGAAA	797
QY	277	nGlySerCyseGlyAspProGlyLleProAlaTYArgArgGluGlySerArgPheHl	297
DB	798	GGAGAGGTGGGGATCTCGAATCCCCGCTATGGAAAGGAGGAGGACAGATGTTCTCT	857
QY	297	SHISGLYASRTHrLeuLYSRPheGluCYseGlnProAlaPheGluLeuValLGLYGLNLYSRAl	317
DB	858	CCATGGAGATACACTCACCTTGAATGCCCGCGCTTTGAGCTGGTGGGGAGAGAT	917
QY	317	aLleThrCyseGlnLYSRASnaAnGlnTTPSerAlaLeuLYSRProGlyCyseValPheSerCy	337
DB	918	TATCACCTGTGACGAACATCATGTGCTGGCAACAAGCCAGCTGATTTTCATG	977
QY	337	sPhePheAsnPheThrSerProSeGlyValValLeuSerProAntLYRProGluAspTY	357
DB	978	TTTCTTCAACTTATGAGCATCATCTGGGATTAATCTGTGCACCAATATTCCAGAGATA	1037
QY	357	rGlyAsnHlVleuHlSCYseValTPLeuLleLeuAlaArgProGluSerArgLleHle	377
DB	1038	TGGGAACAACATGAACCTGTGTCTGTTATTATCTCGAGGACGAGAGATTCAAATTCACCT	1097
QY	377	uAlaPheAsnAspLleAspValGluProGlnPheAspPheLeuValLleLYSRGlyAl	397
DB	1098	AATCTTAAATGATTTGATGATGAGACCTCAATTTACCTTCCGCGGCACAGAGATGG	1157
QY	397	aThrAlaGluAlaProValLeuGlyThrPheSerGlyAsnGlnLeuProSerSerLleTh	417
DB	1158	CATTTCGACAACTACTGTCTGGGTACTTTTCTCGCAATGAAGCTTCCACAGCTGGC	1217
QY	417	rSerSerGlyHlVleAlaArgLeuGlnPheGlnThrAspHlSerThrGlyLYSRGly	437
DB	1218	CAGCAGTGGGCAATAGTTGCTTGGAATTTCACTGACCATTCACACTGCGACAGAG	1277
QY	437	YPheAsnLleThrPheThrThrPheArgHlAsnGluCYSRAspAspProGlyValProVa	457
DB	1278	GTTCAACATCACTTACACCACTTTGGTCAGATAGAGGCCATGATCTCGCATTCAT	1337
QY	457	lAsnGlyLYSRPheGlyAspSerLeuGlnLeuGlySerSerLleSerPheLeuCYAs	477
DB	1338	AAACGAGACGACTTTTGGTGAAGATTTCTACTCGGAGCTCGTTTCTTCCACTGGA	1397
QY	477	pGluGlyPheLeuGlyThrGlnGlySerGluThrLleThrCYValLeuLYSRGluLYSR	497
DB	1398	TGATGGCTTTTGCAAGACCCAGGAGATCCGAGTCCATTCCTGCATATCGCAAGCGGGA	1457
QY	497	rValValTPheAsnSerAlaValLeuArgCYseGlnLYSRProCYseGlyGlyHlVleuThSe	517
DB	1458	CGTGGTCTGAGACTCACCGTCCCCCGCTGTGAACCTCATGTGGTGAACATTCGACGC	1517
QY	517	rProSerGlyThrLleLeuSerProGlyLYRProGlyPheThrLYSRAspAlaLeuSerCY	537
DB	1518	GTCACGCGGAGCATTTTGCTCCCTCGATGCGCAGAGATATTATTAAGATTCCTTACATG	1577
QY	537	aAlaTPValleGluAlaGlnProGlyLYRProLleLYSRHlPheAspArgPheLY	557

Db	1578	TCATGTGATATATGAGCAAAACAGGCGACTCTATCAAAATTAATTCTTTGACAGATTTC	1637
Qy	557	STHGIuValAsnTyraSPThrLeuGIuValaIarAspGIyArGThrTySerAlaProLe	577
Db	1638	GACAGAGGTCAATTATATGACACTTGGAGGTGACAGATGCGGCACAGTTCGCCCACT	1697
Qy	577	uILEGIyValTyriSGIyThrGIuValProGIuPheLeuILESerThSerAnTyrLe	597
Db	1698	GATCGGCAGATGACACGCGCACCCAGGACCCCACTTCCTCATTCAGCACCGGGAATTCA	1757
Qy	597	uTYrLeuLeuPheSerThraSPlySerHisSerAspIIGIyPheGIuLeuArGyTgI	617
Db	1758	GTACTGCTATATCACCACTGACACAGCGCTCCAGCATCGGCTTCCTCATTCAC	1817
Qy	617	uThrILEThrLeuGIuInSerAspHisCySLeuAspProGIyILEProValAsnGIuIar	637
Db	1818	GAGGTGACGCTTAGTGAGTTCCTGCTCGACCCGGGACATCCCTTGAAACGGCACATCG	1877
Qy	637	gHISGIyAsnAspPheTyTValaIaLeuValThrPheSerCyAspAspGIyTyTh	657
Db	1878	CCACGGTGGAGACTTGGCATCAGGTGCACAGTACTTCAGCTGTGACCCGGGGTAC	1937
Qy	657	rLeuSerAspGIyIuPProLeuGIuCySGIuPProAsnPheGIuThrPseArGIaIaLeuP	677
Db	1938	ACTAAGTACACAGAGCCCTGCTGTGAGAGGACACCAAGTGAACACAGCCTTGGCC	1997
Qy	677	oSeCySLeuIaLeuCySGIyGIyPheILEGIySerSerGIyThrILELeuSerP	697
Db	1998	CAGCTGCAGACCTCTATGTGAGGCTACTCCAAAGGAAGTGGAAACAGTCTTCTCC	2057
Qy	697	oGIyPheProAspPheTyTProAsnAsnLeuAsnCySThrTrpILEIGIuThrSerH	717
Db	2058	TGGGTTTCCAGATTTTATCCAACTCTCTAACTGCACCGTGACCATTTGAAGTGTCTCA	2117
Qy	717	SGIyTySGIyValPhePheThrPheHISThrPheHISLeuGIuSerGIyHISAspTyTLe	737
Db	2118	TGGGAAAGAGATTCAAAATGATCTTTCACACTTCATCTTGAAGATTCACACGACTATT	2177
Qy	737	uLeuILEThrGIuAsnGIySerPheThnGIuPProLeuHrGIuInuThrGIySerArGLe	757
Db	2178	ACTGATCCACAGAGATGAAATTTTCCGAGCCCGTTCACAGCTCACCGGCTGGGTCT	2237
Qy	757	uProAlaProILESerAlaGIyLeuTyTGIyAsnPheThraIaGIuValaIarGheILESe	777
Db	2238	GCCTCATACGATCAAGGACGCGCTGTTGGAAACTTCACCTCCAGCATTCGGTTTATATC	2297
Qy	777	rAspPheSerMetSerTyTGIuGIyPheAsnILEThrPheSerGIuTyTAspLeuGIuP	797
Db	2298	AGACTTCCAAATTTCTGACGAGGCTTCMAATTCAAATTTCCAAATAATGACCTGGAGCC	2357
Qy	797	oCySLeuIuPProGIuValaProAlaTyTSerILEArGIySGIyLeuGIuInPheGIyValaGI	817
Db	2358	ATGTGATATCTTGAAGTCTCGCTCCCTCAGCCGGAAGATTTGGTTTCACTTTGGATGGG	2417
Qy	817	yAspThrLeuThrPheSerCySPhProGIyTyTArGIuLeuGIuGIyThraIaArGIeTh	837
Db	2418	AGACTCTGACGTTTCTGCTTCCTCGGGATATCGTTTAAAGTGCACCAAGCTTAC	2477
Qy	837	rCySLeuGIyGIyArGATgATgLeuThrPseSerPProLeuPProArGySValaIaGIuCy	857
Db	2478	CTGCTGGGTGGGGCGCCGCTGTGAGAGGACCTTCGCAAGGTGTGGCGGAATG	2537
Qy	857	SGIyAsnSerValThrGIyThrGIuGIyThrLeuLeuSerProAsnPheProValAsnTy	877
Db	2538	TGGAGCAAGTGTCAAAGGAATGAAGGAACCTTCTGTCTCCAAATTTTCCATCCAATTA	2597
Qy	877	rAsnAsnAsnHISGIyCySILETyTSerILEIGIuThrGIuPProGIyTySGIyILEGIuLe	897
Db	2598	TGATATATTAACATGAGTGTATCTATAAATTAAGAAACGAAGCGGCAAGGCATCCACT	2657
Qy	897	uLySAlaArgAlaPheGIuLeuSerGIuGIyAspValLeuLySValTyTAspGIyAsnAs	917

Db 2658 TAGAACAGAACTTCCAGTGTGTTGAAGAGATACCTAAAGATATATGATGGAAGA 2717
QY nAsnSerAlaArgLeuLeuGlyValPheSerHisSerGluMetMetGlyValThrLeuAs 937
Db 2718 CAGTTCCTCAGTCACCTGAGGAGCAGCTTCACTAAAGAACTTCCTGGGCTGATCCCTAA 2777
QY nSerThrSerSerSerLeuTrpLeuAspPheIleThrAspAlaGluAsnThrSerLeuVal 957
Db 2778 CAGACATCCCAATACACTGCTGGCTAGAGATTCAACCAATGAGATCTGACACCCAGCAAGG 2837
QY yPheGluLeuHisPheSerSerPheGluLeuIleLeuCyseGlyuAspProGlyThrProly 977
Db 2838 TTTTCACTCACCTCACTCAGTTTGAATCGTGTAAATGATGAGATCCGGGCATCCCTAA 2897
QY sPheGlyThrLeuValHisAspGlyuHisPheAlaGlySerSerValSerPheSerCy 997
Db 2898 CTAGGCTATGAGATCCGTATGAGAGGCACCTTACCGACACTGTATTTCTGACATG 2957
QY sAspProGlyTyzSerLeuArgGlySerGluGluLeuCyseLeuSerGlyGluArgAr 1017
Db 2958 CAACCCGGGGTAGCCCATGATGACGACCAACCCCTACCTGTTTGAAGTGAAGACAGAG 3017
QY gThrTrpAspArgProLeuProThrCyseValAlaGlyCyseGlyThrValArgGly 1037
Db 3018 AGTGTGGAGCAAACTACCTCTGTCATAGCGGAATGTGTGTCATCATCGACG 3077
QY uValSerGlyValLeuSerProGlyTyzProAlaProTyzGluHisAsnLeuAsn 1057
Db 3078 CACATCGAGCAAAATATGTCCTGCTGATCCAGCTCCGTATGACAAACCTCCAC 3137
QY sIleTrpThrIleGluAlaGluAlaGlyCyseThrIleGlyLeuHisPheLeuValPheAs 1077
Db 3138 CACCTGGATTATGAGGACAGACCCAGGAAGACCATTAAGCTCATTTCTTTTCCA 3197
QY pThrGluGluValHisAspValLeuArgIleTrpAspGlyProValGluSerGlyVal 1097
Db 3198 CACGGAATGCTCACGACATCTCTCAAGGTCTGGAGCGGCGGTGACAGTACATCT 3257
QY uLeuTyzGluLeuSerGlyProAlaLeuProTyzAspLeuHisSerThrPheAsnSer 1117
Db 3258 GCTGAAGAGTGAAGTGTCCGCCCTTCCGAGAGACATCACAGACCTTCAACTAC 3317
QY lValLeuGlnPheSerThrAspPheThrSerLeuGlyPheAlaIleGlnPheSe 1137
Db 3318 CACCTGCACTTCGACAGCACTTCTTCATCAGCAAGTGTGCTTCATCATCACTTC 3377
QY rValSerThrAlaThrSerCyseAsnAspProGlyIleProGlnAsnGlySerArgSer 1157
Db 3378 CACCTCAATTGACGCCACTGTACACATCCAGGTATGCCCAAAATGGGACCCGCTATG 3437
QY yAspSerTrpGluAlaGlyAspSerThrValPheGlnCyseAspProGlyTyzAlaLeu 1177
Db 3438 AGACACAGAGAGGCTGAGACACCGTCACATTCAGGTGACCCCTGGCTATAGCTCCA 3497
QY nGlySerAlaGluIleSerCyseValIleGluAsnArgPheThrProGlnProSer 1197
Db 3498 AGGACAAAGCCAAATACCTGTGTGACGATTAACCGGCTTTTGGCAACCAAGCC 3557
QY oProThrCyseIleAlaProCyseGlyuHisPheLeuThrGlyProSerGlyValIle 1217
Db 3558 TCTTACATGATAGCTGCTGTGTGAGAGATCTGACGGGCCCAAGGAGTATTTTGTGC 3617
QY rProAsnTyzProGluProTyzProProGlyuHisCyseAspTrpLeuValThrVal 1237
Db 3618 ACCCACTACCCACACCGATCTCTCTGGAAAGAAATGTGACTGAGAGTAAAGTGA 3677
QY rProAspTyzValIleAlaLeuValPheAsnIlePheAsnLeuGluProGlyTyzAsp 1257
Db 3678 CCCGGACTTGTATGCTCTGTATATCAAAAGTTTCAACATGAGACCAGCTATGACT 3737
QY eLeuHisIleTyzAspGlyArgAspSerLeuSerProLeuIleGlySerPheTyz 1277
Db 3738 CTTACACATCTATGAAAGGGAAGATTCCAACAGCCCTTCATTGGAGTTTACAGGGCTC 3797

QY rGlnLeuProGlyValArgIleGluSerSerAsnSerLeuPheLeuAlaPheArgSer 1277
Db 3798 TCAGGCCCCAGAAAGAAATGAGAGTAGCCGAAACAGCCTGTTTGGCATTTCCGAGTGA 3857
QY pAlaSerValSerAsnAlaGlyPheValIleAspTyzThrArgValGlySerAspLeuLeu 1297
Db 3858 TGCTTCCTGGGCTTTCAGGGGTTGCCCATTTAAATTAAGAAACCAAGGGAAGCTTG 3917
QY sPheAspProGlySerIleLeuAsnGlyThrArgValGlySerAspLeuLeuGly 1317
Db 3918 TTTTGACCCAGGAAATATATGAAATGGGACAAAGATGGGAACAGACTTCAAGCTTGCTC 3977
QY rSerValThrTyzTyzCyseHisGlyTyzGluValGluGlyThrSerThrLeuSer 1337
Db 3978 CACATCACCTACAGTGTGACTGTGGCTATATCTTTGACCCCTCATCATCACCTG 4037
QY sIleLeuGlyProAspGlyTyzProValTyzAsnAspProArgProValCyseThrAla 1357
Db 4038 TGTGATGGGGCTGATGGGAAACCTCTCGGAGCAACAGCTGCTCCCTCGAAATGTCC 4097
QY oCyseGlyGlyGlnTyzValGlySerAspGlyValValLeuSerProAsnTyzProGln 1397
Db 4098 CTGTGGAGGCTAGTACAGGATCAGAGGGGTAGTTTATCACCAACTACCCCATTA 4157
QY nTyzThrSerGlyGlnIleCyseLeuTyzPheValThrValProTyzAspTyzValAla 1417
Db 4158 TTACACAGCTGTCAAAATATGCTCTATTCATCAGGTACCGTAACAAAGAAATTTGTGCTT 4217
QY eGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspValValGluValHisAspGly 1437
Db 4218 TGGACAGTTTGGCTATTTCCAGACAGCCCTGAATATTTGGAGAAATTTATGATGGAAC 4277
QY sSerGlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThrGlyuSerLeu 1457
Db 4278 CCATCAGAGGCCAGACTTCTCAGCTCACTCGGGGTCTCACTCAGGAGGAAACATTGCC 4337
QY oLeuAlaThrSerAsnGlnValLeuIleLeuPheSerAlaIleGlyLeuAlaProAla 1477
Db 4338 CTGGCTACGTCAAAATCAAAATTTGCTGCTCAGATCGTCAAGAGCGGTGCTTGGCCG 4397
QY gGlyPheHisPheValTyzGlnAlaValProArgThrSerAlaThrGlnCyseSer 1497
Db 4398 CGGCTTCCACTTCGTATCAAGCTGTTCTCGTACAGTACACCCATATGAGCTGCT 4457
QY lProGluProArgTyzGlyTyzArgLeuGlySerAspPheSerValGlyAlaIleVal 1517
Db 4458 CCCCAGGCCAGATACGGAAGAGAAATGTGTGAGTTTCTGCCGCTCATGCTCG 4517
QY gPheGlyCyseAsnSerGlyTyzAlaLeuGlnGlySerProGluIleGlyuCyseLeuPro 1537
Db 4518 ATTGCAAGTCAACCCGGGATACCTGCTCAGAGTTTCCAGCGGCTCCACTGCACTGCT 4577
QY lProGlyAlaLeuAlaGlnTyzAsnValSerAlaProThrCyseValProCyseGly 1557
Db 4578 GCCCAAGCCTTTGGACAGTGAAGACACAGATCCCAAGCTGTGTGATACCTCGAGG 4637
QY yAsnLeuThrGluArgArgGlyThrIleLeuSerProGlyPheProGluProTyzLeu 1577
Db 4638 CAATTTCACTCAAGCAAGATACATATCTCTCCCGGCTACCTGAGCATACGGA 4697
QY nSerLeuAsnCyseValTrpTyzIleValProGluGlyAlaGlyIleGlnIleGln 1597
Db 4698 CACTTAAGCTGTATATGAAATCATATGTTTACGAGGGCTCGGGAATTCAGATCCA 4757
QY lValSerPheValThrGluGlnAsnTyzAspSerLeuGluValPheAspGlyAlaAsp 1617
Db 4758 GATCAGTTTCCACAGGAGCAAGACTGGACTCCCTTGAGATCCACATGTGTGGAGAT 4817
QY nThrValThrMetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSer 1637
Db 4818 GACCAGCCAGACTGGAGAACTTCTCAGGACCAACAGTACCGGACACTGTGAACAGTAC 4877

Qy 1637 rSerAnGlnLeuTyrLeuHisPheTyrSerAspIleSerValSerAlaIleGlnPheH 1657
Db 4878 TTCAACCAACTACTACCTGCATTTCCAGTCTGACATTAGTGTGACAGTGTGTTTCCA 4937
Qy 1657 sLeuGlnTyrIleThrValGlyLeuSerSerCysProGlnProAlaValProSerAnG 1677
Db 4938 CCTGGAATACAAACTGAGTCTTGTGTCGATGCCAAGAACACCCCTCCACCAACAG 4997
Qy 1677 yValIleThrValGlnLysTyrLeuValAsnAspValValSerPheGlnCysGlnProG 1697
Db 4998 CATCAAAATCGAATCGTACATGAGTGAACAGACGTGCTCCCTCCAGTCCGAGCCCG 5057
Qy 1697 yTyrAlaLeuGlnGlyHisAlaHisIleSerCysMetProGlyThrValAlaGAGTTPAs 1717
Db 5058 GTACACCTCGAGGCGCTTCCACATTCTCTGATGCCAGGACCGTTCGCTTGGA 5117
Qy 1717 nTyrProProLeuLeuCysIleAlaGlnCysGlyValThrValGlnLysMetGlnGlyVa 1737
Db 5118 CTATCCGTCCTCCCTGTCGATTCGACACTGTGAGGAGCGCTGAGCACCTGGGTGGTGT 5177
Qy 1737 lIleLeuSerProGlyPheProGlyAsnTyrProSerAsnMetCysSerTTPylsI 1757
Db 5178 GATCCTGAGCCCGCGCTTCCAGGTTCTTACCCCAACAACTTAGACCTGACCTGAGAGAT 5237
Qy 1757 eAlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheserThrGluProAs 1777
Db 5238 CTCATTACCCATCGGCTGATGTCACATATTCACTTTCTGATTTTCTTACCGAAGCTAA 5297
Qy 1777 nHisAspTyrIleGlnIleArgAnGlyProTyrGlnThrSerArgMetGlyArgPh 1797
Db 5298 TCATGACTTCTCTGAAATTCAAAATGAGCTTACACACCGCCCATGATGAGACATT 5357
Qy 1797 eSerGlySerGluLeuProSerSerLeuLeuSerThrSerHISgluThrValTyrPh 1817
Db 5358 TAGGGACCGGATCTCCCGCGCTGTCGAGCACACGATGAAACCTCATCCACTT 5417
Qy 1817 eHisSerAspHisSerGlnAsnArgProGlyPheLeuLeuGlnTyrGlnAlaTyrGlnLe 1837
Db 5418 TTATAGTACCATTCGCAAAACCGGCAAGATTAACTTCTTACCAAGTATGAGACA 5477
Qy 1837 uGlnGluCysProAspProGlu 1844
Db 5478 ACAACGAGAACCGAACCACAA 5499

RESULT 9
US-10-276-934-4
Sequence 4, Application US/10276934
Publication No. US20030180750A1
GENERAL INFORMATION:
APPLICANT: University of Leeds
APPLICANT: Matham, Alexander F.
APPLICANT: Jackson, Andrew P.
APPLICANT: Woods, Christopher G.
TITLE OF INVENTION: Treatment of Cancer and Neurological Diseases
FILE REFERENCE: 9052-144
CURRENT APPLICATION NUMBER: US/10/276,934
CURRENT FILING DATE: 2002-11-20
PRIORITY APPLICATION NUMBER: PCT/GB01/02240
PRIORITY FILING DATE: 2001-05-21
PRIORITY APPLICATION NUMBER: GB0012186.3
PRIORITY FILING DATE: 2000-05-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 5667
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc_feature
LOCATION: (588)..(588)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc_feature

LOCATION: (624)..(624)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (672)..(672)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1265)..(1265)
OTHER INFORMATION: "n" is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (5641)..(5641)
OTHER INFORMATION: "n" is any nucleotide
US-10-276-934-4
Alignment Scores:
Pred. No.: 0
Score: 6978.00
Percent Similarity: 81.12%
Best Local Similarity: 67.03%
Query Match: 41.08%
DB: 15
Gaps: 7
US-10-016-248-2 (1-3104) x US-10-276-934-4 (1-5667)
Qy 36 ValIleYsGlnIleGlnLeuLysSerArgGlyValIleYsLeuMetProSerIleAspAsn 55
Db 66 GTGMAAAAGCGATGTTGATGAGTCAAGAGTCAAGATGCTGCCACAGAGATGA 125
Qy 56 SerGlnLysThrSerVal 61
Db 126 AGCCATMAAACTGTGCTGTCGATATGAAATTAACATATACCATGTACATCCAGAGAGA 185
Qy 62 -----LeuThrGln-ValGlyValSerGlnG 70
Db 186 TAGCTTATTATGATGTCGGMAAAGCATGCTTCAAGTAAATTCAGTGAAGCCAAAG 245
Qy 70 Y-----HisAsnMetCysProAspProGlyIleProGluArgGlyLysArg 85
Db 246 AGGTGTTCATTTGCTGTCATGATGTCATGATGTCATGATGTCATGATGTCATGATG 305
Qy 85 gLeuGlySerAspPheArgLeuGlySerSerValGlnPheThrCysAnGlnGlyTyrAs 105
Db 306 AGCAGGTTCCGACTTCCAGGTTGTCGAAATGTAACATTTTTCATGATGAGCAATTGCT 365
Qy 105 pLeuGlnGlySerIleAspArgIleThrCysMetLysValSerAspMetPheAlaIleTTPSe 125
Db 366 GCTTCAGGAGTCTTAAAGCATTCACCTCTGACAGAGTTCACAGAGCGCTGCTTGAG 425
Qy 125 rAspHisArgProValCysArgAlaArgMetCysAspAlaHisLeuArgGlyProSerG 145
Db 426 TGACACAGGCGCATCTGCGAGCGAGAACATGTGATGCATCAATCTGCTGGGCCACAGCG 485
Qy 145 yIleIleThrSerProAsnPherProIleGlnTyrAspAlaAsnAlaHisCysValTTP 165
Db 486 CGTATTAATCTCCCTTAATATCCGTTCCAGTATGAAGAAATGACACATGTTGTGGGT 545
Qy 165 eIleThrAlaLeuAsnProSerIleValIleYsLeuAlaPheGlnGluPheAspLeuG 185
Db 546 CATCACACACACGACCGGACCAAGGTCATCAAGCTTGTTCATGAGAGATTGAGCTGGA 605
Qy 185 uArgGlyTyrAspThrIleuThrValGlyAspArgGlyGlnAspArgIleAspGlnYsTh 205
Db 606 GCGAGGCTATGACACCTTACCGTGGTGTGATGCTGGAGAGTGGAGACACCGATGCTGT 665
Qy 205 lLeuTyrMetSerGlnAsnAlaCysSerAspSerProHisThrProGlySerArgIlePr 225
Db 666 CTGTGTANGT-----CTACGAGGATTCAGTGTTC 695
Qy 225 ogIu-----SerMetSerGlyAspIleTTPArgGlnLysTyrThrValIleuGlnI 242
Db 696 TGACCTTATGTTGATGATGAGCAACCAAGATGTGGTACAT----- 735

QY 242 eCyArGAspIleSerSerSerAspAlaArgSerGlySerValArgIysSerProIySth 262
Db 736 -----CTGAGTCGAGTGAATAGC----- 753
QY 262 rSerAsnAlaValGluLeuValAlaProGly-----ThrgIuileGluG1 277
Db 754 -----ATTGGCTCAGCTGGGTTAAAGCTGTTAAAGCTTTACCAAGAAATGAAAA 797
QY 277 ngIySerCyGglYAspProGlyIleProAlaTyGlyArgArgGluGlySerArgPheH1 297
Db 798 GGGAGGGGTGGGATCTCGAATCCCGCTATGGGAAGGAGCGGAGCACTTTCTCT 857
QY 297 sH1sGlyAspThrLeuIySPhegIuCyGlnProAlaPheGluLeuValGlyGlnIySAl 317
Db 858 CCAAGAGATACATCTACCTTGATGATCCCGCGGCTTTGAGTGGTGGGAGAGACT 917
QY 317 aIleThrCySglnIySAsnAengInTPSerAlaIySAspProGlyCySValPheSerCy 337
Db 918 TATCACCCTGTCAGAGAAACATCAGTGGTGGCAACAGCCAGCTGTGATTTTCATG 977
QY 337 sPhePheAsnPherTherSerProSerGlyValValLeuSerProAsnTyProGluAspTy 357
Db 978 TTCTTCACTTACGGGATCATCTGGGATTAATCTGTCAACAAATTAATCCAGAGATA 1037
QY 357 rGlyAsnH1sleuH1sCySValTPLeuIleLeuAlaArgProGluSerArgIleH1sle 377
Db 1038 TGGGAAACAACATGAACGTGTCTGTGATTAATCTCGAGGACCGAAGTGAATTCACCT 1097
QY 377 uAlaPheAsnAspIleAspValGluProGlnPheAspPheLeuValIleYsAspGlyAl 397
Db 1098 AATCTTTAATGATTTGATGTGAGCTCAATTGATTTGACCTTCTCGGGTCAAGATGTGG 1157
QY 397 aThrAlaGluAlaProValleuGlyThrPheSerGlyAsnGlnLeuProSerSerIleH1 417
Db 1158 CATTTCTGACATACTGTCGGGTACTTTTCTGGGAATGAAGTGCCTTCCAGCTGGC 1217
QY 417 rSerSerGlyH1sValAlaArgLeuGluPheGlnThrAspH1sSerThrgIySAspGly 437
Db 1218 CAGCAGTGGGATATAGTTCGCTTGGAAATTCAGTGCACATTCACATCTGGCAGAG 1277
QY 437 yPheAsnIleThrPheThrThrPheArgH1sAsnGlyCySProAspProGlyValProVa 457
Db 1278 GTTCACATCATCTACACCAATTTGTCAGAAATGATGTCATGATCTCGCATTCCTAT 1337
QY 457 lAsnGlyLyAspPheGlyAspSerLeuGlnLeuGlySerSerIleSerPheLeuCySAs 477
Db 1338 AAACGGACGAGCTTTTGTGACAGCTTCTAATCGGAGGCTCGGTTCTTTCACATGTA 1397
QY 477 pGluGlyPheLeuGlyThrgInIySerGluThrIleThrCySValLeuIySglIySAs 497
Db 1398 TGATGGCTTGTCAAGACCAGGATCCGATCGATCACTGCAATCTGCAACACGGAA 1457
QY 497 rValValATTPAsnSerAlaValLeuArgCySglIyAlaProCyGlyGlyH1sleuThrSe 517
Db 1458 CGTGGTCTGAGCTCCACCGTCCGCTGTGAAGCTCCATGTGTGGGACATCTGACAG 1517
QY 517 rProSerGlyThrIleLeuSerProGlyTyProIleIySleThrPheAspArgPheIy 537
Db 1518 GTCCAGGAGGATTTTGGCTCTGATGCGAGATTAATTAAGGATTTCTTAATG 1577
QY 537 sAlaATPValIleGluAlaGlnProGlyTyProIleIySleThrPheAspArgPheIy 557
Db 1578 TGAATGATATTAAGAACAAACACAGGCCACTCTATCAAAATTAATCTTGAAGATTTGA 1637
QY 557 sThrgIyValAsnTyAspThrLeuGluValArgAspGlyArgThrTySerAlaProIe 577
Db 1638 GAAAGAGTAATTAATGACCTTGAAGTCAAGATGGCCACAGCTTGTGCCACACT 1697
QY 577 uIleGlyValTyH1sGlyThrgInValProGlnPheLeuIleSerThrSerAsnTyIle 597
Db 1698 GATCGGAGATACACAGGACCCAGGACCCAGTTCTTCATACGACCGGAGACTTCAT 1757
QY 597 uTyrLeuLeuPheSerThrAspLySerH1sSerAspIleGlyPheGlnLeuArgTyG1 617
Db 1758 GATACCTGATTAATCACATGACAAACGCCCTCCAGACATCGGCTTCATCAATATGA 1817
QY 617 uThrIleThrLeuGlnSerAspH1sCySLeuAspProGlyIleProValAsnGlyGlnAr 637
Db 1818 GAGTGTAGCGCTTGATCGGATTCCTGCTGAGACCGGCGCATCTCTGTAAACGCGCATG 1877
QY 637 gH1sGlyAsnAspPheTyValAlaLeuValThrPheSerCySAspSerGlyTyTh 657
Db 1878 CCAAGGTGAGACTTTGGGATCAGGTCCACAGTACTTTCACTGTGACCCGGGATACAC 1937
QY 657 rLeuSerAspGlyGluProLeuGluCySglIyProAsnPhenGlnTPSerArgAlaLeuPr 677
Db 1938 ACTAATGACAGAGCCCTGCTGTGTAGAGAAACCAAGTGAACACAGCTTGGC 1997
QY 677 oSerCySgluAlaLeuCySglYgIyPheIleGlnIySerSerGlyThrIleLeuSerPr 697
Db 1998 CAGCTGCAAGCTCTATGTGAGGCTACATCCAAAGGAAGATGGAAACAGTCTTCTCC 2057
QY 697 oGlyPheProAspPheTyProAsnAsnLeuAsnCySthrTPIleIleGluThrSerH1 717
Db 2058 TGGGTTTCAAGATTTTATCCAAACTCTTAACCTGACGTGAGCCATTTGAAGTGTCTCA 2117
QY 717 sGlyLySgIyValPhePheThrPheH1sleuGlnIySerGlyH1sAspTyIle 737
Db 2118 TGGGAAGAGATTCAAATGATCTTTCACACCTTATCTTGAGATTTCCACAGCATATT 2177
QY 737 uLeuIleThrgIuAengIySerPheThrgInProLeuArgInLeuThrgIySerArgIe 757
Db 2178 ACTGATACACAGAGATGGAAGTTTTCGAGCCGCTGACAGCTCACGGGTCCGGTGT 2237
QY 757 uProAlaProIleSerAlaGlyLeuTyGlyAsnPherH1sleuValArgPheIleSe 777
Db 2238 GCCATACAGATCAAGGACGCTGTGTGGAACCTTCACTGCCAGCTTCGCTTATATATC 2297
QY 777 rAspPheSerMetSerTyGluGlyPheAsnIleThrPheSerGlyTyAspLeuGluPr 797
Db 2298 AGACTTCTCAATTTGTGAGAGGCTTCAATATCAATTTTCAGAAATATGACTTGAGACC 2357
QY 797 oCySgluGluProGluValProAlaTySerIleArgLySgIyLeuGlnPheGlyValG1 817
Db 2358 ATGTATGATCTGTGAGTCCCGCTTCAGCCGGAAGAAATTTGGTTTCACTTGTGTGG 2417
QY 817 yAspThrLeuThrPheSerCySPherProGlyTyThrgInIySerGlyThrAlaArgIleH1 837
Db 2418 AGACTCTGAGGTTTCTGCTTCTCGGATGTGTTGAAGGTGCCACCAAGCTTAC 2477
QY 837 rCySleuGlyGlyArgArgArgLeuTySerSerProLeuProArgCySValAlaGlyCy 857
Db 2478 CTGCTGGGTGGGGGCCCGCGTGTGTGAGATGACCTCTGCCAAGTGTGTGGCGAATG 2537
QY 857 sGlyAsnSerValThrgIyThrgInGlyThrLeuLeuSerProAsnPherProValAsnTy 877
Db 2538 TGGACCAAGTGCAAAGGAATAAGAAACATTTACTGTCCTCAAAATTTTCCATCAATTA 2597
QY 877 rAsnAsnAsnH1sGlyCySgIleTySerIleGlnThrgInProGlyIySgIyIleGlnIe 897
Db 2598 TGAATATACCATGTGTATATCTTAATAATGAACAAACCCGCAAGGGGATCCACT 2657
QY 897 uLySAlaArgAlaPheGluLeuSerGluGlyAspValLeuIySValTyAspGlyAsnAs 917
Db 2658 TAGAACACAGAGCTTCCAGCTGTTTGAAGAGATATCTTAAGATATGATGGAAGAA 2717
QY 917 nAsnSerAlaArgLeuLeuGlyValPheSerH1sSerGluMetMetGlyValThrLeuAs 937
Db 2718 CAGTTCCTCACGATCCAGGCGACGTTCACTTAATAATTAATCTTGGGGCTGATCTTAA 2777
QY 937 nSerThrSerSerSerLeuTyProLeuAspPheIleThrAspAlaGluAsnThrSerLyG1 957
Db 2778 CAGCAGATCCATACCTGTGTGTAGATTGAACCAATGATCTGACACGCAAGG 2837
QY 957 yPheGluLeuH1sPheSerSerPheGluLeuIleYsCySgluAspProGlyThrProLy 977

```
Db 2838 TTTTCACTCACCATACCACTTTTGTGATCTGTAAATGTAGATTCGGGCATCCCTAA 2897
Qy 977 sPhgGlyTyIlyValHisAepGluGlyHisPheAlaGlySerValSerPheSerCy 997
Db 2898 CTACGGCTATGTAGATCCCTGTATGAAGGCCATTACCGACATCTGATTCGTACAGTTG 2957
Qy 997 sAepProGlyTySerLeuArgGlySerGluGluLeuLeuCyLeuSerGlyGlyValArg 1017
Db 2958 CAACCCGGGGTACCGCATGCAATGGCAGACACCTGTACCTGTTGTAGTGAAGACAGAG 3017
Qy 1017 gThrTrpAspArgProLeuProThrCyValAlaGlyGlyGlyThrValAlaGlyGly 1037
Db 3018 AGTGTGGAGCAACAACACTACTCTGTGATAGCCGAATGTGTGTGATGATCCATGACAC 3077
Qy 1037 uValSerGlyGlnValLeuSerProGlyTyTrpAlaProTyGluHisAenLeuAenCy 1057
Db 3078 CACATTCAGAGCAAGAAATATGTCCTCGCTATCCAGCTCCGTATGACAAACATCCACTG 3137
Qy 1057 sIleTrpThrIleGluAlaGlyValGlyCySerThrIleGlyLeuHisPheLeuValPheAs 1077
Db 3138 CACCTGGATTATAGAGGAGCAACCAAGAAAGACATTAAGCTCCATTCATTGTTTTCGA 3197
Qy 1077 pThrGluGluValHisAepValLeuArgIleTrpAspGlyProValGlySerGlyValle 1097
Db 3198 CACGGAGATGGCTCAGCAGCATCTCTAAGCTGTGGGACGGCGGTGAGACATGACATCTC 3257
Qy 1097 uLeuGlyGluLeuSerGlyProAlaLeuProIlyAepLeuHisSerThrPheAenSerVa 1117
Db 3258 GCTAGAGAGTGGAGTGGCTCCGGCTCCGAGAGACATCCACAGCACCCTTCACATCACTCA 3317
Qy 1117 lValLeuGlnPheSerThrAepPhePheThrSerIlyGlnGlyPheAlaIleGlnPheSe 1137
Db 3318 CACCTGGAGATTCAACAGCGACTTCTTATCAGCAAGTGTGCTTCATCCATCCAGTTCTC 3377
Qy 1137 rValSerThrAlaThrSerCyAsnAepProGlyIleProGlnAenGlySerAArgSerGly 1157
Db 3378 CACCTCAATTGCAACGACACTCTTAACGATCCAGTATGCCCAAAATGGCACCCCTATGCG 3437
Qy 1157 yAepSerTrpGluAlaGlyAepSerThrValPheGlnCyAepProGlyTyTralAleuGly 1177
Db 3438 AGACAGAGAGAGGCTGGAGACACCGTCACATTCACAGTGTACCTGGCTATCAGCTCCA 3497
Qy 1177 nGlySerAlaGluIleSerCyValIlyLeuGluAenAepPhePheTrpGlnProSerPr 1197
Db 3498 AGGCAAGCCAAATATCACCTGTGTGACGTGAATTAACCGGTTCTTTGGCAACAGACC 3557
Qy 1197 oProThrCyAlaIleAepProCyGlyGlyAepLeuThrGlyProSerGlyValIleLeuSe 1217
Db 3558 TCCATCACTGATAGCTGCTGTGTGAGGGAATCTGACGGGCCACAGAGTGTATTTTGTCTC 3617
Qy 1217 rProAenTyTrpGluProTyTrpProGlyTyLeuGluCyAepTrpIlyValIleValSe 1237
Db 3618 ACCCAACTACCCACGACCGATCTCTCCGTGGAGAAATGTGATCGAGAGTAAAGTGA 3677
Qy 1237 rProAepTyTrAlIleAlaLeuValPheAenIlePheAenLeuGluProGlyTyTrAepPh 1257
Db 3678 CCCGAGCTTGTGTCTGCTGTGATATTCAAAGTTTCAACATGAGCCCAAGCTATGACTT 3737
Qy 1257 eLeuHisIleTyAepGlyArgAepSerLeuSerProLeuIleGlySerPheTyArgIlySe 1277
Db 3738 CCTACACATCTATGAAGGGAAGATTCACAGCCCTCTCTATGGAGTTACAGGGGCTC 3797
Qy 1277 rGlnLeuProGlyValIleGlySerSerAenSerLeuPheLeuAlaPheArgSerAe 1297
Db 3798 TCAGGCCCCCAAGAAATAGAGATGAGCGAAACAGCTGTTTCTGGCATTTCCGAGTGA 3857
Qy 1297 pAlaSerValSerAenAlaGlyPheValIleAepTyTrhGluAenProAenArgIlySerCy 1317
Db 3858 TGCCTCCGCGGCTTTCAAGGCTTCGCAATTGAATTTAAAGAAACACAGCGAGAGCTTG 3917
Qy 1317 sPheAepProGlySerIleLeuAenGlyThrArgValGlySerAepLeuTyLeuGlySe 1337
Db 3918 TTTTGACCCAGAAATATATGAATGGGACAGAGATTGGAACAGACTTCAAGCTTGGCTC 3977
Qy 1337 rSerValThrTyTrCySerHisGlyGlyTyTrGluValGluGlyThrSerThrLeuSerCy 1357
Db 3978 CACCATCAGCTTACCAAGGTGATCTCTGGCTATTAATTTTGAACCCCTATCCATCCACTCG 4037
Qy 1357 sIleLeuGlyProAepGlyLysProValTTPAsnAenProAepProValCyThrAlaPr 1377
Db 4038 TGTATTTGGGCTGATGGAAACCTCTGGAGCAAAAGTGTGCTGCTCTGCAATGCTCC 4097
Qy 1377 oCyGlyGlyGlnTyValGlySerAepGlyValIleLeuSerProAenTyTrProGlnAs 1397
Db 4098 CTGTGAGGCAAGTACACGGGATCAGAAAGGGGTATTTTATACCAAACTAACCCCATTA 4157
Qy 1397 nTyTrhSerGlyGlnIleCySerLeuTyPheValThrValProIlyAepTyTrAlaPhe 1417
Db 4158 TTTACACGCTGTGCAAAATATAGCTCTTATCCATCAGCGTACCAAGAAATTCGTGCTTT 4217
Qy 1417 eGlyGlnPheAlaPhePheHisThrAlaLeuAenAepValValGlyValHisAepGlyHis 1437
Db 4218 TGGACAGTTTGCTTATTTCCAGACAGCCCTGAATGATTTGGCAGAAATTTATTTGAGAAC 4277
Qy 1437 sSerGlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThrGlyGlySerLeuPr 1457
Db 4278 CCATGCAACAGGCCAGACTTCTCAGCTCAGCTCGGGGTCTTCATCAGGGGAAACATTTGCC 4337
Qy 1457 oLeuAlaThrSerAenGlnValLeuIleTyPheSerAlaIlyGlyLeuAlaProAlaAr 1477
Db 4338 CTGGCTACACGTCAAAATATTCCTCCATTCAGTGAAGAGCGGTGCTCTGCGCG 4397
Qy 1477 gGlyPheHisPheValTyTrGlnAlaValProArgThrSerAlaThrGlnCySerSerVa 1497
Db 4398 CGGCTTCACTTCCTGTATCAAGCTGTCTCTGATCCAGTGAACCCCAATGCACTCTGT 4457
Qy 1497 lProGluProArgTyGlyIlyArgLeuGlySerAepPheSerValGlyAlaIleValAr 1517
Db 4458 CCGGAGCCCAAGATACGGAAGAGAAATGGTTCGATTTTCTCGCGGCTCATCGTCCG 4517
Qy 1517 gPheGluCyAenSerGlyTyTrAlaLeuGlnGlySerProGluIleGluCySerLeuProVa 1537
Db 4518 ATTGAGTGTGAACCCGGGATATCCTGTACAGGGTTCACAGCGGCTCCACGCTCCAGTCCGT 4577
Qy 1537 lProGlyAlaLeuAlaGlnTrpAenValSerAlaProThrCyValValProCyGlyGly 1557
Db 4578 GCCCAAGCCTTGGCAAGTGAACGACAGATCCCAAGCTGTGTGTATCCCTGCACTG 4637
Qy 1557 yAenLeuThrGluArgArgIlyThrIleLeuSerProGlyPheProGluProTyTrLeuAs 1577
Db 4638 CAATTTCACTCAACGAGAGAGTACATCTGTCCCGGGCTACCTGACGCATACGGA 4697
Qy 1577 nSerLeuAenCyValTrpIlyValIleValProGluGlyAlaGlyIleGlnIleGlnVa 1597
Db 4698 CAACTGAACTGATATGAGAAATCATAGTTACGAGAGGGCTCGGGAATTCAGATCCAAAT 4757
Qy 1597 lValSerPheValThrGluGlnAenTrpAepSerLeuGluValPheAepGlyAlaAepAs 1617
Db 4758 GATAGTTTGGCCAGGAGCAAGCTGGAGCTCCTTGAGATTCACAGATGTGGGAGTGT 4817
Qy 1617 nThrValThrMetLeuGlySerPheSerGlyThrThrValProAlaLeuAenAenSerTh 1637
Db 4818 GACCGCAACCGAGCTGGGAAGCTTCTCAGGCAACAGAGTACCGGCACTGGCGAAACATAC 4877
Qy 1637 rSerAenGlnLeuTyTrLeuHisPheTyTrSerAepIleSerValSerAlaAlaGlyPheHis 1657
Db 4878 TTTCAACCAACTACCTGATTTCCAGTCTGACATTAAGTGTGACACTCTGTTTCCA 4937
Qy 1657 sLeuGluTyTrLeuThrValGlyLeuSerSerCyProGluProAlaIlyProSerAenGly 1677
Db 4938 CTTGGAATTAACAAACTGATAGCTTCTGCTGATGCCAAGAACAGCCCTCCCGACACAG 4997
Qy 1677 yValTySerThrGlyGluArgTyTrLeuValAenAepValAenSerPheGlnCyGlyProGly 1697
Db 4998 CATCAAAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5057
```

```

QY 1697 YTYALaleuGlnGlyHisAlaHisSerCysMetProGlyThrValArgArgTPas 1717
DB 5058 GTACACCTTCGACGAGCGCTTCCACATTTCTGTATGCCAGGACCGTTCGCGCTTGGA 5117
QY 1717 nTYrProProLeuGlyHisAlaGlnCysGlyGlyThrValGlnGluMetGlnGlyVal 1737
DB 5118 CTATCCGCTCCCTGTCATTCGACCTGTCGAGGAGCGCTGACGACCTTGCGTGTGT 5177
QY 1737 lIleLeuSerProGlyPheProGlyAsnTYrProSerAspMetCysSerTPlyS11 1757
DB 5178 GATCCGAGCCCGGAGCTTCCAGGATCTTACCCCAACACTTAACGACCTGAGAGAT 5237
QY 1757 eAlaLeuProValGlyPheGlyAlaHisSerGlnPheLeuAsnPheserThrGluProAs 1777
DB 5238 CTGATTCACCATCGCGCATGTCATATTCATTCATTTCTGAATTTTCTACCGAGCTAA 5297
QY 1777 nHisAspTYrIleGluIleArgAngGlyProTYrGluThrSerArgMetGlyArgPh 1797
DB 5298 TCATGACCTTCCTGAAATTCAAATGACCTTACACACAGCCCATGATGACAACT 5357
QY 1797 eSerGlySerGluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTYrPh 1817
DB 5358 TAGGGACGAGATCTCCCGGCGCTGTCGAGCAACGATGAAACCTTCATTCACCTT 5417
QY 1817 eHisSerAspHisSerGlnAsnArgProGlyPheLeuGlnGlyValGlnAla 1834
DB 5418 TTATAGTACCATTCGCAAAACCGGCAAGATTTAACTTGCTTACCAAGCC 5469

RESULT 10
US-10-467-042-27
; Sequence 27, Application US/10467042
; Publication No. US20040077048A1
; GENERAL INFORMATION:
; APPLICANT: WARREN, Bridget A.; HONCHELL, Cynthia D.;
; APPLICANT: LU, Yan; CHAWLA, Narinder K.;
; APPLICANT: BURFORD, Neil; DELESEANE, Angelo M.;
; APPLICANT: GANDHI, Ameena R.; BAUDHN, Mariah R.;
; APPLICANT: GRIFFIN, Jennifer A.; GIERZEN, Kimberly J.;
; APPLICANT: LU, Dzung Alma M.; ISON, Craig H.;
; APPLICANT: RAMKUMAR, Jayalakshmi; TANG, Y. Tom;
; APPLICANT: LAL, Preeti G.; BOROMSKY, Mark L.;
; APPLICANT: DUGGAN, Brendan M.; HAPALIA, April J.A.;
; APPLICANT: ARVIZU, Chandra S.; TRANNGAVELU, Kavitha;
; APPLICANT: VAO, Monique G.; ELLIOTT, Vicki S.;
; APPLICANT: DING, Li; YUE, Henry;
; APPLICANT: LEE, Sally; SWARNAKAR, Anita;
; APPLICANT: TRAN, Uyen K.; XU, Yuning
; TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES
; FILE REFERENCE: PI-0361 USN
; CURRENT APPLICATION NUMBER: US/10/467,042
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: PCT/US02/02813
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 60/265,705
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/266,762
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/269,581
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/271,198
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/272,813
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: US 60/275,586
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: US 60/278,505
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/280,539
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 27

```

```

; LENGTH: 4506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7473634CB1
US-10-467-042-27

Alignment Scores:
Pred. No: 0
Score: 4148.50
Percent Similarity: 71.53%
Best Local Similarity: 57.45%
Query Match: 24.42%
DB: 16 Gaps: 15

US-10-016-248-2 (1-3104) x US-10-467-042-27 (1-4506)
QY 1 MetAlaGlyAlaProProAlaLeuLeu-----LeuProCysSerLeu 15
DB 385 ATCAGTCGACGAGAGAGCTTCATTCATATGCTAATCGCATGACCTCCCTCTCCAGTT 444
QY 16 lIleSerAsp-----CysCysAlaSerAsnGlnArgHisSer 27
DB 445 ATCAGTACGACAGAAATTGGCTACACTTCATTCACCTTGACAGCAACCGACGCAAA 504
QY 28 ValGlyValGlyProSerGluLeuVallylylyGlnIleGluLeuLYSerArgGlyVal 47
DB 505 ---GGATTTAAAGCTTCAGTTCCAAAGTAAAGCGGATTAAGTGAAGTCAAGAGAGTGC 561
QY 48 lYsLeuMetProSerLYsAspAsnSerGlnLYsThrSerValLeuThrGlnValGlyVal 67
DB 562 AGATGCTGCCAGACAGAGATGAGAGCCATTAATACTGCTTGAAGCCAGAGAGTGT 621
QY 68 SerGlnGlyHisAspMetCysProAspProGlyIleProGlnArgGlyLYsArgLeuGly 87
DB 622 GGATTCGCTTCCTGACATGTGTCACATCTCGGATTCAGAAATGTGAAGACAGAGT 681
QY 88 SerAspPheArgLeuGlySerSerValGlnPheThrCysAsnGlnGlyTYrAspLeuGln 107
DB 682 TCCGATTCAGGATGGTGGTGCATATGACAGTTCATGAGACACATTAATGCTCCAG 741
QY 108 GlySerLYsArgIleThrCysMetLYsValSerAspMetPheAlaAlaTrpSerAspHis 127
DB 742 GGATCTAAAGATCATCCTGACAGAGATTAAGAGAGCGCTGCTTGAGTACCA 801
QY 128 ArgProValCysArgAlaArgMetCysAspAlaHisLeuArgGlyProSerGlyIleIle 147
DB 802 AGGCCCATCTGCCGAGCGAGAACATGATGATCAATCTGCGGCGCCAGCGCGTCAT 861
QY 148 ThrSerProAspPheProIleGlnTYrAspAsnAlaHisCysValTYrIleIleThr 167
DB 862 ACCTCCCTTAATTAATCCGTTAGTATGAAGTAAATGACACTGTGTGTGGTCAATCAC 921
QY 168 AlaLeuAsnProSerLYsValIleLYsLeuAlaPheGlnGluPheAspLeuGlnArgGly 187
DB 922 ACCACGACCGCGAGACAGATCATCAAGTTCCTTGAAGAGTTGAGCTGAGGAGGAGC 981
QY 188 TYrAspTYrLeuThrValGlyAspGlyGlnAspGlyAspGlnLYsThrValLeuTYr 207
DB 982 TATGACACCTGACGATGGTGTGATCTCGGAGAGGAGGAGACACGATCGCTCTTGATC 1041
QY 208 MetSerGlnAsnAlaCysSerAspSerProHisThrProGlySerArgIleProGlu--- 226
DB 1042 GTG-----CTCAGCGGATCCAGTTCCTACCTC 1071
QY 227 -----SerMetSerGlyAspIleTrpArgGlnLYsTrpThrValLeuGlnIleCysArg 244
DB 1072 ATTGACATGACAGCAACGATGTGCTTAC----- 1104
QY 245 AspIleSerSerSerAspAlaArgSerGlySerValArgLYsSerProLYsThrSerAsn 264
DB 1105 ---CTGACGTGGATGATAGC----- 1122

```

Qy 265 AlavAlgluLeuValAlaProGly-----ThrGluIleGluGlnGlySer 279
::: :::::
Db 1123 -----ATTGGCTCAGCTGGGTTTAAAGCTGTTTACCAAGAAATTGAAAGGAGG 1173
Qy 280 CysGlyAspProGlyIleProAlaTyrGlyArgGluGlySerArgPheHisIleGly 299
::: :::::
Db 1174 TGTGGGGATCTGGGAATCCCGGCTATGGGAAGCGGAGCAGAGTTTCTCATGGA 1233
Qy 300 AspThrLeuLysPheGluCysGlnProAlaPheGluLeuValGlyGlnLysAlaIleThr 319
1234 GATCACTGACCTTGAATGCCGCGCGCTTTGAGCGTGGGGAGAGATTATCAAC 1293
Qy 320 CysGlnLysAsnAsnGlnTyrSerAlaLysLysProGlyCysValPheSerCysPhePhe 339
1294 TGTACGACGAACAAATCACTGATGCTGGCAACAAAGCCAGCTGTATTTTCAATGTTTCTTC 1353
Qy 340 AsnPheThrSerProSerGlyValValLeuSerProAsnTyrProGluAspTyrGlyAsn 359
1354 AACTTACGGCATCATCTGGGATTATCTGTACCAAAATTATCCAGAGAAATATGGGAAAC 1413
Qy 360 HisLeuHisCysValTyrPheLysIleLeuAlaArgProGluSerArgIleHisLeuAlaPhe 379
1414 AACATGAACGTGTCTGGTGTATATCTCGAGACGAGAGATCGAATTCACCTAATCTTT 1473
Qy 380 AsnAspIleAspValGluProGlnPheAspPheLeuValIleLysAspGlyAlaIleThrAla 399
1474 AATATTTTGTATGTTGACCTCAATTTGACTTCTCGCGTCAAGAGATGATGCAATTTCT 1533
Qy 400 GluAlaProValLeuGlyThrPheSerGlyAsnGlnLeuProSerSerIleThrSerSer 419
1534 GACATACTGTCCTGGGACTTTTCTCGCAATGAAAGCTTCCAGCTGGCAGAGAGCT 1593
Qy 420 GlyHisValAlaArgLeuGluPheGlnThrAspHisSerThrGlyLysArgGlyPheAsn 439
1594 GGGCATATAGTTCCTTGGAAATTCAGATCAGATTCACATTCAGTGGAGAGGATTCAC 1653
Qy 440 IleThrPheThrThrPheArgHisAsnGluCysProAspProGlyValProValAsnGly 459
1654 ATCTCTTACACCAATTTGGTCAAAATGAGTGCATGATCTGGGATTCCTATTAACGGA 1713
Db 460 LysArgPheGlyAspSerLeuGlnLeuGlySerSerIleSerPheLeuCysAspGluGly 479
1714 CGAGTTTGTGGTGAAGGTTTCTAATCGGAGAGCTCGGTTCTTCCAGTGTGATGATGCG 1773
Qy 480 PheLeuGlyThrGlnGlySerGluThrIleThrCysValLeuLysGluGlySerValVal 499
1774 TTTGTCAAGAACCCGAGGATCCGAGTCCATTACTGCAATACGCAAGACGGGAAAGTGGTC 1833
Qy 500 TrpAsnSerAlaValLeuArgCysGluAlaProCysGlyGlyHisIleuThrSerProSer 519
1834 TGGAGCTCCACCGGCCCCGCTGTGAGCTCATGTGGTGAATCTGACAGCTCCAGC 1893
Qy 520 GlyThrIleLeuSerProGlyTyrProGlyPheTyrLysAspAlaLeuSerCysAlaIleTyr 539
1894 GGAATCATTTTGGCTCTCGGATGCGAGATATTAAGATTCCTTTACATGTGATGCG 1953
Qy 540 ValIleGluAlaGlnProGlyTyrProIleLysIleThrPheAspArgPheLysThrGlu 559
1954 ATAATTAGCAAAACAGGCGCTATCAAAATACTTTGACAGATTTCAAGACAGAG 2013
Qy 560 ValAsnTyrAspThrLysGluValArgAspGlyArgThrTyrSerAlaProLeuIleGly 579
2014 GTCAATATATGACACTTGGAGTGAAGATGAGGCGCAGCCAGTGTGCTCCCATATGCGC 2073
Qy 580 ValTyrHisGlyThrGlnValProGlnPheLeuIleSerThrSerAsnTyrLeuTyrLeu 599
2074 GAGATACACGCGCACCGGACCCAGTCTCTATACGACACCGGGAATTCATTAATCTG 2133
Qy 600 LeuPheSerThrAspLysSerHisSerAspIleGlyPheGlnLeuAspTyrGluThrIle 619
2134 CTGTTTCAACACTGACAAACAGCGCTCCAGCATCGGTTCTCTCACTCACTAATGAGATGTC 2193

Qy 620 ThrLeuGlnSerAspHisCysLeuAspProGlyIleProValAsnGlyGlnArgHisIleGly 639
2194 ACGCTTAGAGTGGATTCCTGCTGAGACCCGGGATTCCTGTGAACGGGCAATCGCACCGCT 2253
Qy 640 AsnAspPheTyrValGlyAlaLeuValThrPheSerCysAspSerGlyTyrThrLeuSer 659
2254 GGAATCATTTGGCATCAGGTCCACAGTCACTTCACTGTGACCCGGGGTACACATAAGT 2313
Qy 660 AspGlyLysProLeuGluCysGluProAsnPheGlnTyrPheSerArgAlaLeuProSerCys 679
2314 GACACAGAGCCCTGCTGTGAGAGGAAACACAGAGGAAACAGCGCTTGGCCAGGCTC 2373
Qy 680 GluAlaLeuCysGlyGlyPheIleGlnGlySerSerGlyThrIleLeuSerProGlyPhe 699
2374 GACCTCTATGTGAGGCTCATCTCAAGGAAAGTGAACAGTCTTCTCTCGGCTT 2433
Qy 700 ProAspPheTyrProAsnAsnLeuAsnCysThrTyrIleIleGluThrSerHisGlyLys 719
2434 CCAGATTTTATCCAAACTCTTAACCTGACAGTGCACATGAAAGTGTCTCATGGGAA 2493
Qy 720 GlyValPhePheThrPheHisIleThrPheHisLeuGlnSerGlyHisAspTyrLeuLeuIle 739
2494 GGAATTCAAATGATCTTTCACACCTTTCATCTTGAAGTTCACGACTATTTACTGATC 2553
Qy 740 ThrGluAsnGlySerPheThrGlnProLeuArgGlnLeuThrGlySerArgLeuProAla 759
2554 ACAGAGATGAGATTTTTCGAGCCGCTTGCAGGCTCACCGGATGGGTGCTTCAT 2613
Qy 760 ProIleSerAlaGlyLeuTyrGlyAsnPheThrAlaGlnValArgPheIleSerAspPhe 779
2614 ACGATCAAGGAGGCTGTTTGGAACTTCACTGACCTCCAGCTCGGTTATATCAGACTTC 2673
Qy 780 SerMetSerTyrGluGlyPheAsnIleThrPheSerGluTyrAspLeuGluProCysGlu 799
2674 TCAATTTTCGACGAGGCTTCAATATCAATTTTCAGAAATATGACCTGAGGCCATGTGAT 2733
Qy 800 GluProGluValProAlaTyrSerIleArgLysGlyLeuGlnPheGlyValGlyAspThr 819
2734 GATCTGAGATCTCTGCTTCAAGCAAAAGTGTTCATCTTGGTGTGGAGACTCT 2793
Qy 820 LeuThrPheSerCysPheProGlyTyrArgLeuGluGlyThrAlaArgIleThrCysLeu 839
2794 CTGACGTTTCTCTGCTTCTCGGAGTATCGTTTAAAGTGCACACAGCTTACTGCTG 2853
Qy 840 GlyLysArgArgLeuTyrPheSerProLeuProArgCysValAlaGluCysGlyAsn 859
2854 GGTGGGGCGCGCTGTGTGAGTGCACCTCTGCAAGGTGTGTGGCGGAATGAGCA 2913
Qy 860 SerValThrGlyThrGlnGlyThrLeuLeuSerProAsnPheProValAsnTyrAsnAsn 879
2914 AGTGTCAAGGAATGAAGAACATTTCTGTCCAAATTTTCCATTCATTAATGATAT 2973
Qy 880 AsnHisGluCysIleTyrSerIleGlnThrGlnProGlyLysGlyIleGlnLeuLysAla 899
2974 AACCATAGAGTATCTTAATAATGAAACAGGACCGGACGATCCACCTTAGAACA 3033
Qy 900 ArgAlaPheGluLeuSerGluGlyAspValLeuLysValTyrAspGlyAsnAsnAsnSer 919
3034 CGAAGCTTCCAGCTGTGTGAAGGATACCTAAGGTATATGATGAAAAAGACAGTTC 3093
Qy 920 AlaArgLeuLeuGlyValPheSerHisSerGluMetThrGlyValThrLeuAsnSerThr 939
3094 TCAGTCCAGTGGGACAGTTCACATAAAATGAATCTTGGGCTGATCTTAACAGCA 3153
Qy 940 SerSerSerLeuTyrPheAspPheIleThrAspAlaGluAsnThrSerGlyPheGlu 959
3154 TCAATACACTGTGGCTAGAGTTCAACACCAATGAGATGACACCGACCAAGTTTCAA 3213
Qy 960 LeuHisPheSerSerPheGluLeuIleLysCysGluAspProGlyTyrProLysPheGly 979
3214 CTCACTTATACAGTGTGTGATCTGTGTAATAATGTGAGATCCGGGCAATCCCTAATCGCG 3273
Qy 980 TyrLysValHisAspGluGlnHisPheAlaGlySerSerValSerPheSerCysAspPro 999

```

Db 3274 TATAGATCCGTGATGAAGGACACTTATCCGACACTGATGTTCTGATCAGTTGACACCG 3333
Qy 1000 G1YrSerLeuArgS1SerGluGluLeuLeuCysLeuSerG1YglValArgThrThr 1019
Db 3334 GGGTACCGCATGATGATGACACACCTGACCTGTTGATGAGAGAGAGTGG 3393
Qy 1020 AspArgProLeuProThrCysValAlaGluCysG1YglYThrValArgG1YglValSer 1039
Db 3394 GACAAACCACTACCTTGTGATGACCGAATGTGGTGTGATGATCCATGACGACATCA 3453
Qy 1040 G1YglValLeuSerProG1YrProAlaProTyrGluHisLeuLeuSerCys1eTyr 1059
Db 3454 GGAGCAATATGTGCTGCTGATCCAGCTCCGATGACAAACCTCCACTGACCTGG 3513
Qy 1060 Thr1leGluAlaG1YglCysThr1leG1YleuHisPheLeuValPheAspThrGlu 1079
Db 3514 ATTATAGGCGACGCCGAAAGACCTTAGCCTCCATTTGATTTTTCACACGAG 3573
Qy 1080 GluValHisAspValLeuArg1eTyrAspG1YProValG1YSerG1YValLeuLys 1099
Db 3574 ATGCTCAGACGACATCTTCAAGCTCTGGGACGGCGGATGACATGATCTGCTGAG 3633
Qy 1100 GluLeuSerG1YProAlaLeuProLysAspLeuHisSerThrPheAsnSerValValLeu 1119
Db 3634 GAGTGGAGTGGCTCCGCTTCGAGAGACATCCACAGACCTTCACTCACTCACCTG 3693
Qy 1120 GluPheSerThrAspPhePheThrSer1YglGln1YleuAla1leGlnPheSerValSer 1139
Db 3694 CAGTTCCAGACGACCTTCTCATACGACAGCTGCTTCTCCATCCAGTTCTCCAGATCT 3753
Qy 1140 ThrAlaThrSer-----CysAsnAspProG1YleProGlnAsnG1Y----- 1153
Db 3754 CAGCTGGAACACCAAGACGCTGCTGACCAACCCAGCTCATTGACGATCTCTC 3813
Qy 1154 -----SerArgSerG1YAspSer1YrGluAlaG1YAspSerThrValPhe 1168
Db 3814 CACAAATGTAGCTTGGCCACTTCTGGAACCAAGTGAATCGGACACATCTCCATCTC 3873
Qy 1168 ----- 1168
Db 3874 CCTGAAACCTGATAAATTGACTCTCTACCTGCAATGAAATCAATGCTGTTTCT 3933
Qy 1169 -----GlnCysAspProG1YrAlaLeuGlnG1YSerAlaGlu1e 1182
Db 3934 AGTTTGAAGAAAGAGGTTCTTAATACCTCAGTCTGATTAAGAAACGACCCAGTTA 3993
Qy 1183 SerCysValLys1e-GluAsnArgPhePheThrGlnProSerProProThrCys1leAl 1202
Db 3994 CCTGCTTCACTGACGAAAGAACTGGGCTGTTATGTCTCTCACTCCACCAATTCG 4053
Qy 1202 aProCysG1YAsp-----LeuThrG1YProse 1212
Db 4054 TCCCTCAGTGGGAAATCCAGCATGAAATCAATCAAGCTGATCTTCCAAACCAAA 4113
Qy 1212 rG1YVal1leLeuSerProAsn1YrProGluProTyrProProG1YLysGlu----- 1229
Db 4114 GTGGGAAACTCTTCAAAAGTCAAAACAGCTGTCATCACACCAAGAAAGCATCACT 4173
Qy 1230 -----CysAsp1YrPheValThrVal-SerProAsp1YrVal1leAlaLeuValPheAl 1247
Db 4174 ACTCTTTTGTAGTGGAGATGGGATGGGATGACATGCACTCTCACTTTTGTCTTA 4233
Qy 1247 sn1lePheAsnLeuGluProG1Y-----TyrAspPheLeuHis1leYrAspG1YA 1264
Db 4234 AT-----GAACTGCGACCACAAAATATGAGCATTAACATATAGATGATG 4281
Qy 1264 rGAspSerLeuSerProLeu1e 1271
Db 4282 TGTGCTTCATTAACCTGCTAATT 4304

```

```

; Sequence 3, Application US/09799514
; Patent No. US20020065220A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: Immunoglobulin Superfamily Polynucleotides, Polypeptides, and A
; FILE REFERENCE: PTO15P1
; CURRENT APPLICATION NUMBER: US/09/799,514
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/US00/23662
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/152,248
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-514-3
Alignment Scores:
Pred. No.: 0 Length: 2487
Score: 3942.00 Matches: 711
Percent Similarity: 99.03% Conservative: 1
Best Local Similarity: 98.89% Mismatches: 7
Query Match: 23.21% Indels: 0
DB: 9 Gaps: 0
US-10-016-248-2 (1-3104) x US-09-799-514-3 (1-2487)
Qy 2080 AspSerThrG1YVal1leLeuSerGlnSer1YrProG1YSer1YrProGlnPheGlnThr 2099
Db 1 GACTCAGACGGGCTGATCTCTGACGACAGCTACCTCGAAGCTATCCAGTTCCAGAC 60
Qy 2100 CysSer1YrLeuValArgValGluProAsp1YrAsn1leSerLeuThrValGluTyrPhe 2119
Db 61 TGTCTTGGCTGTGATGAGTGGACCGCACTTAACATCTCCCTCACTGAGTACTTC 120
Qy 2120 LeuSerGluLysG1YrAspG1YlePheAspG1YProSerG1YGlnSerPro 2139
Db 121 CTCAGGAAACCAATATGATGATTTGATTTGATGCTTCATCAGACAGAGTCTT 180
Qy 2140 LeuLeuValAlaLeuSerG1YAsn1YrSerAlaProLeu1leValThrSerSerAsn 2159
Db 181 CTGCGAAAGCCCTCAGTGGAAATTAATCAAGCTCCCTGATGTCACAGCTCAAGCAC 240
Qy 2160 SerVal1YrLeuArgTyrSerSerAspHisAla1YrAsnArgG1YPheLys1leArg 2179
Db 241 TCTGTGACTGTGCTGTGATGATCAAGCTTAATCGGAAGGCTTCAAGATCCGC 300
Qy 2180 TyrSerAlaProTyrCysSerLeuProArgAlaProLeuHisG1YPhe1leLeuGln 2199
Db 301 TATTAGACCCCTTACTGACGCTGCGGAGGCTCCATCTCCATGCTTCAATCTTAAGCCAG 360
Qy 2200 ThrSerThrGlnProG1YG1YSer1leHisPheG1YCysAsnAlaG1YrArgLeuVal 2219
Db 361 ACCAGACCCAGCCCGGGGCTCCATCACTTGGCTCAACAGCGGCTACCGCTGGTG 420
Qy 2220 G1YHisSerMetAla1leCysThrArgHisProGlnG1YrHisLeuThrSerGlnAla 2239
Db 421 GGACACAGATGAGCATCTGTACCGGACCCCGGCTCAACACTGTGAGCAAAACC 480
Qy 2240 IleProLeuCysGlnAlaLeuSerCysG1YleuProG1YAlaProLysAsnG1YMetVal 2259
Db 481 ATCCCTCTCTGTCMAAGCTCTTTCGTGTGGCTCTCTGAGGCCCCCAAAATGAAATGGTG 540
Qy 2260 PheG1YLysGluTyrThrValG1YrThrValAlaVal1YrSerCysSerGluG1YrHis 2279
Db 541 TTGGCAAGAGATACACAGTGGAAACCAAGCCGCTGACACTGACAGTGAAGGCTACAC 600
Qy 2280 LeuGlnAlaG1YAlaGluAlaThrAlaGluCysLeuAspThrG1YLeuTyrSerAsnArg 2299
Db 601 CTCAGGAGGCGCTGAAGCCACTGCAAGTGTCTGACACAGGCTTAATGAGCAACCGC 660

```



```

; PRIOR APPLICATION NUMBER: US 60/255,622
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 31
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7473630CBI
US-10-399-455-31

Alignment Scores:
Pred. No.: 4,4e-238 Length: 2609
Score: 2694.00 Matches: 493
Percent Similarity: 73.17% Conservative: 126
Best Local Similarity: 58.27% Mismatches: 175
Query Match: 15.86% Indels: 54
DB: Gaps: 4

US-10-016-248-2 (1-3104) x US-10-399-455-31 (1-2609)

QY 887 ILeGlnThrGlnProGlyLysGlyIleGlnLeuLysAlaArgAlaPheGluLeuSerGlu 906
DB 124 GTCAGGAGCTCCAGCAATCAGGAAATCTGACTCCAGGAGTGAAGAG-----CAG 177
QY 907 GlysPheValLeuLysValIleAspGlyAspAsnAsnSerAlaArgLeuLeuGlyValPhe 926
DB 178 GGCACATGGC-AGAGATTATGATGAGAAAGATMAAGCATCATCTACTAGAGTGTCTTT 236
QY 927 SerHisSerGluMetMetGlyValThrLeuAsnSerThrSerSerSerLeuThrPLeuAsp 946
DB 237 ACTGGTCATCTATGCGCGAGCTGACACTTAGTAGTCTTCAATCAACTCTGGCTTGA 296
QY 947 PheIleThrAspAlaGluAsnThrSerLysGlyPheGluLeu----- 960
DB 297 TTTAATTCGATACGTGAAGGAGCAGATGAAGGCTTTCACCTGTGTATACCAAGAGATC 356
QY 960 ----- 960
DB 357 ATAGTATTGACAGAGAGCTACAGTACTACATTGACAGATACAGAACAGAGCAGAA 416
QY 961 -----HisPheSerSerPheGluLeuIleLysCysGluAspProGlyThrPro 976
DB 417 CATTTGCTAGGAGAG-TCAGGTTTGAATCTGCACACTGTGAAGTCTCTGGCATTTCA 475
QY 977 LysPheGlyThrLysValHisAspGluGlyHisPheAlaLysSerSerValSerPheSer 996
DB 476 CAATTGGATACAGATCAGTACAGGACGACTTGTGTTGACACCATTTTATGGA 535
QY 997 CysAspProGlyThrSerLeuArgLysSerGluGluLeuLeuLysSerGlyLysArg 1016
DB 536 TGCATTCAGAGCTACACTCCACGAGAGT----- 565
QY 1017 ArgThrTrpAspArgProLeuProThrCysVal-AlaGluCysGlyGlyThrValArg 1036
DB 566 -----AGCCTTCACTGAATGTGAAGTGTGAAGTGTGAAG 601
QY 1036 YgiValSerGlyValLeuSerProGlyThrProAlaProGlyGluHisAsnLeuAs 1056
DB 602 AGAATCTTCAGGAAGATCTTATCTCTGCTATCTCTTCCATATGACAAATAACCTGG 661
QY 1056 nCyValIleThrThrIleGluAlaGluLysCysThrIleGlyLeuHisPheLeuValPh 1076
DB 662 TTGCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
QY 1076 eAspThrGluGluValHisAspValLeuArgIleTrpAspGlyProValGluSerGlyVa 1096
DB 722 TGATAGGAGAGCATCATGATATATCTCGAGTGTGAGAGGCTCCAGAAATATGATAT 781
QY 1096 IleuLeuLysGluLeuSerGlyProAlaLeuProLysAspLeuHisSerThrPheAsn 1116
DB 1116 ----- 1116

```

```

DB 782 GCTTTAAAGAAATTAGTGATCTTTATCTCTGAGAGAAATTCATAGCACTCAATAT 841
QY 1116 rValValLeuGlnPheSerThrAspPheThrSerLysGlnGlyPheAlaIleGlnPh 1136
DB 842 AGTAACCATTCAGATTGACACGAGATTTTATATTACAAATCTGATTTGCAATTCAGTT 901
QY 1136 eSerValSerThrAlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSe 1156
DB 902 TTCAGTCTGTGGCCACTGCGTGTGTGATCCAGAGGAGTCCCATGAAATGGAGCTGGAA 961
QY 1156 rGlyAspSerThrGluAlaGlyAspSerThrValPheGlnCysAspProGlyThrValLe 1176
DB 962 TGGGATGAGAGAAACCTGGGAGACCTGTTTTCATATGATGACCCAGATATAACT 1021
QY 1176 uGlnGlySerAlaGluIleSerCysValLysIleGluAsnArgPhePheThrPLeuProSe 1196
DB 1022 TCAGAGAGAGAGAAAGAAATACCTGATTCAGTAAATCCGTAATCTTGTGACAGCCAG 1081
QY 1196 rProProThrCysIleAlaProCysGlyLysAspLeuThrGlyProSerGlyValIleLe 1216
DB 1082 CCACACAGTCTGTATAGCACCTGTGAGAGCAATTTAACAGGATCTTCAGGCTTATTTCT 1141
QY 1216 uSerProAsnThrProGluProTyProProGlyLysGluCysAspThrLysValThrVa 1236
DB 1142 TTCACCAAACTTCCCTCATTCATATCCGATAGCAGAGCTGTGACTGACTATCACCT 1201
QY 1236 lSerProAspThrValIleAlaLeuValPheAsnIlePheAsnLeuGluProGlyThrAs 1256
DB 1202 CAATGAGACTATGTATATCTCTTGGCGTTCATCAGTTTATACATVAGAACCAACTATGA 1261
QY 1256 PheLeuHisIleLysIleAspGlyArgAspSerLeuSerProLeuIleGlySerPheThrG 1276
DB 1262 CTTCCTCTATATCTATGATGACCCAGACATATAGCCACACTGATTTGAGATTTCAAGA 1321
QY 1276 ySerGlnLeuProGlyArgIleGluSerSerSerAsnSerLeuPheLeuAlaPheArgSe 1296
DB 1322 CAGCAAGTATACAGAGAGAAATGAAAGCAGCTCAATATCAATGATTTGGCTTTGGAG 1381
QY 1296 rAspAlaSerValSerAsnAlaGlyPheValIleAspThrThrGluAsnProArgGluSe 1316
DB 1382 TGATGATCTGTATATTACACTGATTTCACTTCAATTAACAAAGCAAACTCGGAGATC 1441
QY 1316 rCysPheAspProGlySerIleLysAsnGlyThrArgValGlySerAspLeuLysLeuG 1336
DB 1442 CTGCTTTGATCCAGCAATATATGATGACACCACTTGGAAAGCATTTAATTTAGG 1501
QY 1336 ySerSerValThrThrYrCysHisGlyGlyThrGluValGluGlyThrSerThrLeuSe 1356
DB 1502 GTCAACAGCTCACATATCTGATGATGCTGTTATGTTCTTCAAGGTTATTTCACACTCAC 1561
QY 1356 rCysIleLeuGlyProAspGlyLysProValThrAsnAsnProArgProValCysThrAl 1376
DB 1562 CTGTTTCATGGAGATGATGAGAGACCTGAGTGAATGAGACCTTGCACAGTTGTCATGC 1621
QY 1376 aProCysGlyGlyGlnThrValGlySerAspGlyValValLeuSerProAsnThrProG 1396
DB 1622 GCCCTGTGAGAGTGTTCACAGGTTGAGAGGACGCTTTATATCCAAATCATCCAA 1681
QY 1396 nAsnThrThrSerGlyGlnIleCysLeuThrPheValThrValProLysAspThrValVa 1416
DB 1682 AAATTACAGTGTGAGACATTAATGTGTATTTATTCATACAGTTCCAAAGGAGTGTG 1741
QY 1416 lPheGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspValValGluValHisAspG 1436
DB 1742 TTGG---CCAAGTTATTTTTCAGACATCTACACATGTTGTGAGGTGTAGATGG 1798
QY 1436 yHisSerGlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThrGlyGluSerLe 1456
DB 1799 GCCAACTCAGCAATCTTCTGTATCTTCCCTCTCAGAGATCCATTCAGAGAGATCACT 1858
QY 1456 uProLeuAlaThrSerArgGlnValLeuIleLysPheSerAlaLysGlyLeuAlaProAl 1476
DB 1859 TCCACTGAGTTCAAGTATACATGATCAGATTCGATTTCACTTCAAGTGTGACCAATACAGC 1918

```



```

QY 201 AspglnylsThrValleuTyrMetSerGlnAsnAlaCysSerAspSerProHisThrPro 220
Db 601 GACCGAAGACAGATTCTTCAACATC-----CTGACA 630
QY 221 GlySerArgIleProGlnSerMetSerGlyAspIleTyrArgGlnIleTyrThrValleu 240
Db 631 GGTACATCGGTCCGGATCATGTGTGACACCAAT---CATCAAAATGGCTCTCTCTG 687
QY 241 GlnIleCysArgAspIleSerSerSerAspAlaArgSerGlySerValAlaGlyLysSerPro 260
Db 688 CAG-----ACTGATGGC---AGTGGCACT-----708
QY 261 LysThrSerAsnAlaValAlaGlnLeuValAlaProGlyThrGlnIleGlnGlnIleSerCys 280
Db 709 -----TCCCTGGATTCAGGCTTCTTATGAAGATCGAGCGGAGATTGC 756
QY 281 GlyAspProGlyIleProAlaTyrGlyArgArgGlnIleSerArgPheHisArgIleAsp 300
Db 757 GGTACACCTGGCATACCTGCATATGCGCGAGGAGGAGGCTCCGGTTTCAACACGATGAC 816
QY 301 ThrLeuLysPheGlnCysGlnProAlaPheGlnLeuValGlnIleLysAlaIleThrCys 320
Db 817 ACACCTCAAGTTGAGTGCAGCCCGCTTGAGCTGGGAGAGGAGCAATCACAATGC 876
QY 321 GlnLysAsnAsnGlnTyrSerAlaLysLysProGlyCysValPheSerCysPhePheAsn 340
Db 877 CAAMAAGATAACAATGGTGGCTAAGAGAGCGAGCGCTGCTCTCTCTCTCTCTCTCTCT 936
QY 341 PheThrSerProSerGlyValValLeuSerProAsnTyrProGlnAspTyrGlyAsnHis 360
Db 937 TTCAACGAGCCGCTGGGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 996
QY 361 LeuHisCysValTyrPheLeuIleLeuAlaArgProGlnSerArgIleHisLeuAlaPheAsn 380
Db 997 CTCACACTGTGTCTGCTCATCTCTGCGAGGCTGAGAGCGGACATCCATCGGCTTCTTAC 1056
QY 381 AspIleAspValGlnProGlnPheAspPheLeuValIleLysAspGlyAlaThrAlaGln 400
Db 1057 GACATGTACGTGAGGCTCAGATTGATTCTCTGCTCATCAAGATGGGCGGCGCGAG 1116
QY 401 AlaProValLeuGlyThrPheSerGlyAsnGlnLeuProSerSerIleThrSerSerGly 420
Db 1117 GCGCCGCTCTCTGGGACCTTCTCTGAGAAACAGATTCCTCTCTCTCTCTCTCTCTCT 1176
QY 421 HisValAlaArgLeuGlnPheGlnThrAspHisSerThrGlyLysArgGlyPheAsnIle 440
Db 1177 CACGTGCCCCGTCTCGAGTTCAGACTGACCACTCCACAGGAGAGGGGCTTCAACATC 1236
QY 441 ThrPheThrThrPheArgHisAsnGlnCysProAspProGlyValProValAsnGlyLys 460
Db 1237 ACTTTTACCACTTCCGACACAAAGAGTGCCTGGATCTGCTTCAAGTAAATGGCAAA 1296
QY 461 ArgPheGlyAspSerLeuGlnLeuGlnSerSerIleSerPheLeuCysAspGlnIlePhe 480
Db 1297 CGGTTTGGGACAGCTCCAGCTGGGAGCTCCATCTCTCTCTCTCTCTCTCTCTCTCT 1356
QY 481 LeuGlyThrGlnIleSerGlnIleThrIleThrCysValLeuLysGlnIleSerValIleTyr 500
Db 1357 CTGGGAGCTCAGGGCTCAGAGCAATCACTGCTCTTAAGAGAGGCGAGCGTGTCTGG 1416
QY 501 AsnSerAlaValLeuArgCysGlnAlaProCysGlyGlyHisLeuThrSerProSerGly 520
Db 1417 AACAGGCGTGTGTGGGTGTGAAGCTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1476
QY 521 ThrIleLeuSerProGlyTyrProGlyLysPheTyrLysAspAlaLeuSerCysAlaTyrVal 540
Db 1477 ACATCTCTCTCTCGGCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1536
QY 541 IleGlnAlaGlnProGlyTyrProIleLysIleThrPheAspArg 555
Db 1537 ATTGAAGCCAGCCAGGCTTACCCCATCAAAATCACTTGTGACAGG 1581

```

```

RESULT 14
US-10-451-010-20
; Sequence 20, Application US/10451010
; Publication No. US20040082761A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: XU, Yuming
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LEE, Sally
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YUE, Henry
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: HONCHELL, Cynthia D.
; APPLICANT: BURFORD, Neil
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: JACKSON, Jennifer L.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: BANDMAN, Olga
; APPLICANT: GRAUL, Richard C.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: YAO, Monique G.
; APPLICANT: LAL, Preeti G.
; TITLE OF INVENTION: CELL ADHESION PROTEINS
; FILE REFERENCE: PF-0867 USN
; CURRENT APPLICATION NUMBER: US/10/451,010
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/49206
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/256,542
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/259,604
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/260,101
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID NO: 7473626CB1
US-10-451-010-20

Alignment Scores:
Pred. No.: 8.75e-169 Length: 1615
Score: 1940.50 Matches: 386
Percent Similarity: 72.13% Conservative: 10
Best Local Similarity: 70.31% Mismatches: 29
Query Match: 11.42% Indels: 124
DB: 16 Gaps: 7

US-10-016-248-2 (1-3104) x US-10-451-010-20 (1-1615)
QY 1644 HisPheTyrSerAspIleSerValSerAlaIleGlyPheHisLeuGlnIleTyrLys----- 1661
Db 1 CATTTCTACTCAGATATACAGGTATCTGCACTGCTTCCACTTGGAGTACAAAATTTTC 60
QY 1661 ----- 1661
Db 61 TCACCTCAGGCTCAAAACCTTATATCTGCTTCTGCAATCCTAAGTGTCACTACCAA 120
QY 1661 ----- 1661
Db 121 ATACTCTTCAAGAGAAAGATTGTGTGATGTGCTTCTGGGAAACAAGAAAGAGTTTC 180

```

```

1661 ----- 1661
Db 181 TAGGAAACCTGCTGACATCTGGAGCTGCTGCCTACCCCTTTCTCCACTGGGATACAG 240
Qy 1662 -----ThrValGlyLeuSerSerCysProGluProAlaValProSerAsnGlyVal 1678
Db 241 GAAAGAAATAGCGGTGGGCTGAGCAGTGTGCTGGAAACCTGCTGCTGCCAGATGAAGGGGTG 300
Qy 1679 LysThrGlyGluArgTyrLeuValAsnAspValValSerPheGlnCysGluProGlyTyr 1698
Db 301 AAGCTGGCGAGCGCTACTGTGTGATGATGTGTGTCTTTCAGTGTGACCGGGAGAT 360
Qy 1699 AlaLeuGlnGlyValAlaHisIleSerCysMetProGlyThrValArgArgTyrAsnTyr 1718
Db 361 GCCCTCCAGGGGCCAGGCCACATCTCTCGCATGCCGGAACAGTGGCGGATGGAATCAC 420
Qy 1719 ProProProLeuCysIleAlaGlnCysGlyGlyThrValGluGluMetGluGlyValIle 1738
Db 421 CCTCTCCACTCTGATATGACACAGTGTGGGAAACAGTGAAGAGATGAAGGGGTGATC 480
Qy 1739 LeuSerProGlyPheProGlyAsnTyrPheSerAsnMetAspCysSerTyrIleAla 1758
Db 481 CTGAGCCCCGGGCTTCCAGGCACTACCCCACTAATGACATGACTGCTCTGAAATAAGCA 540
Qy 1759 LeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGluProAsnHis 1778
Db 541 CTGCCCGTGGGCTTGGAGCTCACATCCAGTTCCTGAATCTTCCACCGAGCCCAACAC 600
Qy 1779 AspTyrIleGluIleArgAsnGlyProTyrGluThrSerArgMetMetGlyArgPheSer 1798
Db 601 GACTACATAGAAATCCGGATGCGCCCTATGAGACACAGCCCATGATGGAAATTCAGT 660
Qy 1799 GlySerGluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTyrPheHis 1818
Db 661 GGAAGCAGAGCTTCAAGCTCTCTCTCTCCAGTCCACGAGACACACCGTATATTTTCCAC 720
Qy 1819 SerAspHisSerGlnAsnArgProGlyPheLeuGlnGluTyrGlnAlaTyrGluLeuGln 1838
Db 721 AGCAGACCACTCCAGAAATCGGCGCAGGATTCAGCTGAGATTCAGGCTTAAGAACTTCA 780
Qy 1839 GluCysProAspProGluProPheAlaAsnGlyIleValArgGlyAlaGlyTyrAsnVal 1858
Db 781 GAGTGCCAGACCCAGAGCCCTTTGCCAATGGCATTTGAGAGGAGCTGGCTACAACTG 840
Qy 1859 GlyGlnSerValThrPheGluCysLeuProGlyTyrGlnLeuThrGlnHisProValLeu 1878
Db 841 GGACAATCAGAGCTTCAAGTGTGCTGCTCCGGGATCAATGACTGGCCACCTGTCTCTC 900
Qy 1879 ThrCysGlnHisGlyThrAsnArgAsnTyrAspHisProLeuProLysCysGluValPro 1898
Db 901 ACGTGTCAACATGGGACCAACCGGAATCGGAGCAACCCCTGCGCAAGTGAAGTCCCT 960
Qy 1899 CysGlyGlyAsnIleThrSerSerAsnGlyThrValTyrSerProGlyPheProSerPro 1918
Db 961 TGTGGCGGGAACATCACTTCTTCCACGCGCACTGTGTACTCCCGGGGTTCCCTAGCCG 1020
Qy 1919 TyrSerSerSerGlnAspCysValThrPheLeuThrValProIleGlyHisGlyValArg 1938
Db 1021 TACTCCAGCTCCCGAGACTGTGTGCTGATCAACCGTCCCATTTGGCCAGGGGTCCGC 1080
Qy 1939 LeuAsnLeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTyrAspGly 1958
Db 1081 CTCAACTCTCAGCTGCTGCAACAGAGCCCTCTGGAATTTCAATCCATCTGGGATGGG 1140
Qy 1959 ProGlnGlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaValLeuThrVal 1978
Db 1141 CCACAGCAAAACAGACCAACGCGCTCGGCGTCTTCAACCCGAGACATGGCCAAAGAAACAGTG 1200
Qy 1979 GlnSerSerSerAsnGlnValLeuLeuLysPheHisArgAspAlaAlaThrGlyGlyIle 1998
Db 1201 CAGGTTTCATCAACACAGTCTCTGCTCAAGTTTCAACCGTGTATGACCAACGAGGGGATC 1260
Qy 1999 PheAlaIleAlaPheSer-----AlaTyrPro----- 2007

```

```

Db 1261 TTGSCATAGCTTTCTCCGCTCAGTATGAAAGCTGGCTGTGGGAAGGCGAGGCTTT 1320
Qy 2008 -----LeuThrLysCysProProProThrIleLeuProAsn----- 2019
Db 1321 CAAGTCAAGGCTGAGTGTGACTCCGCTGCTCCACATTTGGGAGATGACCTTGAGTG 1380
Qy 2020 -----AlaGluValIleThrGluAsnGluGluPheAsnIleGlyAs 2033
Db 1381 AGTTGTATACCTTTGGAGGCTTCACTGTCTTCAAGATTATGAAATTAATGATTAGC 1440
Qy 2033 pIleValArg-----TyrArgCysLe 2040
Db 1441 CTATGTAGAGCTCTCAGTGCAGGCTTCTGTAAATGCAAGTTTCTCTATTCACACT 1500
Qy 2040 uProGlyPheThrLeuValGlyAsnGluIleLeuThrCysAlaLeuGlyThrTyrLeuG 2060
Db 1501 GCCAGGG-----CAGAGAGGCACAGAAAGCCCA 1527
Qy 2060 nPheGluGlyProProProIleCysGluValHis----- 2071
Db 1528 AACCTTGGTCCAGAGTCCACTATTCACATTCACATCACTGAGTGCATCATCTACT 1587
Qy 2072 -CysProThrAsnGluLeuLeuThr 2079
Db 1588 GTGCCCAAGCTATCCCTTAGCACT 1612

RESULT 15
US-09-911-842-3
; Sequence 3, Application US/09911842
; Patent No. US20020151483A1
; GENERAL INFORMATION:
; APPLICANT: Welcher, Andrew A.
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017/37592
; CURRENT APPLICATION NUMBER: US/09/911,842
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: US 60/222,438
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 11230
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-911-842-3

Alignment Scores:
Pred. No.: 2,296-159 Length: 11230
Score: 1852.00 Matches: 808
Percent Similarity: 31.44% Conservative: 374
Best Local Similarity: 21.49% Mismatches: 1215
Query Match: 10.90% Indels: 1363
DB: 9 Gaps: 150

US-10-016-248-2 (1-3104) x US-09-911-842-3 (1-11230)
Qy 74 CysProAspProGlyIleProGluArgGlyLysArgLeuGlySerAspPheArg----- 91
Db 1310 TGTCTTCCTCCGAGAGCTCCGTAAGAAATGGTTTTTTATACAAACACTTGCAAAACATAC 1369
Qy 92 LeuGlySerSerValGlnPheThrCysAsnGluGlyTyrAspLeuGlnGlySerLysArg 111
Db 1370 TTCATATCCCGCTGTGGGCTCCGATGTGCGCCGGGCTTTGACCTTGGGAGAGCATC 1429
Qy 112 IleThrCysMetLysValSerAspMetPheAlaIleTyrSerAspHisArgProValCys 131
Db 1430 CATTGTGTCAACCC-----AATGCTTTGTGCTGTGGACAGAAAGTTCTGC 1477
Qy 132 ArgAlaArgMetCysAspAlaHisLeuArgGlyProSer---GlyIleIleThrSerPro 150
Db 1478 AGAGTGAAGAGTGC-----CCCACTCCGACAGCCAAACAGGCGACATCAGTGTCTCC 1534

```

QY 151 AsnPro11eg1n1yAspAsnAlaHisCysVal1Trp11e11eThrAlaLeuAsn 170
 Db 1535 ACTGGGAAATGCTCTAC-----AACACCTGTGT-----TTGGTTACCTGC----- 1576
 QY 171 ProSer1yVal11e1yLeuAlaPheGlu1n1PheAspLeuGluA1y1yAspThr 190
 Db 1577 -----AATGAAGATACAGATT 1594
 QY 191 LeuThrVal1G1yAspG1yG1yAspG1yAspG1n1yThrVal1Leu1yTrpMetSerGln 210
 Db 1595 -----GAAGCAGACCTAGCTTACCTGTCAGCA 1624
 QY 211 AsnAlaCysSerAspSerProHis1eThrProG1ySerA1y1eProG1ySerMetSerG1y 230
 Db 1625 AATCCCAAGTGGATGGCCCAAG-----CCCGGTGTGTGA-- 1663
 QY 231 Asp1e1TrpA1yG1n1y1Trp1TrhVal1e1n11e1CysA1yAsp11eSerSerSerAsp 250
 Db 1663 ----- 1663
 QY 251 AlaArgSerG1ySerValA1yG1ySerPro1y1TrhSerAsnAlaVal1G1uLeuValAla 270
 Db 1664 ---CGCATTTGTGCACTTCCAGAACCCCA----- 1693
 QY 271 ProG1y1TrhG1u11eG1uG1n1y1SerCysG1y---AspProG1y11eProAla1yG1y 289
 Db 1694 ---GGCTCATCATTTTCTCCACCCAGCTGGCGCAAGACGCCCGAGGCTT----- 1741
 QY 290 ArgA1yG1uG1ySerA1yPheHis1eG1yAsp1TrhLeu1ySerPheG1yCysG1nProAla 309
 Db 1742 -----GGATGACCTGTGACCTTAAGCTCCCGCAGGA 1774
 QY 310 PheG1uLeuVal1G1yG1n1yAla11eThrCysG1n1yAsnA1y1e1n1TrpSerAla1y1y 329
 Db 1775 TACATTTTATCCGGGGTCAAGAAAGTGAATGTGCCACATCTGGAGTGAAGTGCACAA 1834
 QY 330 ---LysProG1yCysVal1PheSerCysPhe----- 338
 Db 1835 GTTCAGACAGCTGTGTCCAAAGATGTGGAGCTCCCAATCAGCTGTCCAAATGACATT 1894
 QY 339 -----PheAsnPhe1TrhSerProSer--- 345
 Db 1895 GAGGCAAAAGATGGGAGCAGCAGACTGTCTAATGCCACCTGGCAAGTCCCAACAGCT 1954
 QY 346 -----G1yVal1Val1e1ySerProA1y1yProG1yAsp1y1y--- 357
 Db 1955 AAAGCAACTGTGTGAAAAGTGTCT-----AGTCCAGCTCCACCCAGCTTTTACCCCA 2008
 QY 358 -----G1yAsnHis1eHisCysVal1Trp1Leu11eLeuA 369
 Db 2009 CCTTACCTTTCCTCCAAATGGAGACGTGGCATCATCCTACACGGCAACCCAGCTATC----- 2064
 QY 369 1a1yProG1ySerA1yG1y1eHis1eLeuAla-PheAsnAsp11eAspVal1G1uPro----- 386
 Db 2065 ---CGGTAAACCAACCCAGCTGTCTTCTATTAAGTCAATGATGTGAACCGCTGTCT 2122
 QY 387 -----Gln 387
 Db 2123 ATGATTTGGTCCGATCTCCACTCCAAATCAGGTGAGAGAGAGACCTCTGCAAC 2182
 QY 388 PheAspPheLeuVal11e1yAspG1yA1y1a1TrhA1yG1uA1yProVal1e1n1y1TrhPhe 407
 Db 2183 TGGATGAGGCTCAGTTCTCAGACAACTCGGGGCTGAATGGTCTATTACAGCACTAC 2242
 QY 408 Ser---G1yAsnG1n1e1yProSerSer11eThrSerSerG1yHisVal1A1yArgLeuG1n 426
 Db 2243 ACACAAAGGCACTGTTTCTCTCATGGGAAACGGTGTGTGTACACACCC----- 2293
 QY 427 PheG1n1TrhAspHis1e1TrhG1y1yA1yA1yPheAsn11eThrPhe1TrhPheArg 446
 Db 2294 -----ACTGACCCCTCAGGCAACAAGAGCTGTGACATC---CACATTGTCTATAA 2344

QY 447 HisA1yG1uCy1eProAspProG1yVal1ProVal1AsnG1y1yA1yArgPheG1yAspSerLeu 466
 Db 2345 GGTTCCTCCGTGAGGTCCCTTACCCCTGTAAAGGGGACTTTATCTGTGCCAGAT 2404
 QY 467 GlnLeuG1ySerSer11eSerPheLeuCy1eAspG1yG1yPheLeuG1y1TrhG1n1y1ySer 486
 Db 2405 AGTCTGAGATTAACTGTAGCCCTGACCTGCAAGAGGCTTATGATTTACAGAGGCTCA 2464
 QY 487 ---GluThr11eThrCysVal1e1y1yG1uG1y1y-----SerVal1 499
 Db 2465 CCTGAAGATTAAGTCTTGTGAAGATGATCTGAGACACACCATTACTTACAGA 2524
 QY 500 TrpAsnSerAlaVal1e1yA1yCysG1uA1yProCysG1y1y----- 512
 Db 2525 TGGCCAGACTGTGTATAAAGCTTTTGCAAACATGGTTTCAAGTCTTTGAATGCTA 2584
 QY 513 -----G1yHis1e1y1ySerProSerG1y 520
 Db 2585 TACAAACACATCCGTGTGATGACATGATCTGTTAAGAAGTTTCTGACAGATTGAG 2644
 QY 521 Thr11eLeuSerProG1y1yTrpProG1yPhe1y1yAspAla----- 534
 Db 2645 ACTACCTGGGGAACATGTGTCCCTTTTGTAAAGATGATGACATGACATGACAG 2704
 QY 535 -----LeuSerCysAla1TrpVal11eG1uA1yG1nProG1y1yTrpPro----- 548
 Db 2705 CTGAGAGACTGACAAAAAATATGTCATGAGTATTAATTACACTATATAAATGCTTT 2764
 QY 549 -----11e1y11eThrPheAspA1yPhe 556
 Db 2765 GCAATTGACACAGAGGCTGGGTGACAGCAACAGGCTGATTAATCTTCAAGATCACTTC 2824
 QY 557 Lys1TrhG1uVal1Asn1yAsp1TrhLeuG1uVal1A1yAspG1yA1y----- 571
 Db 2825 CTGGATGTTGTACAGAAACACCAACGATGTGGCAAGCCAGATCTGACAGATTAA 2884
 QY 572 -----ThyTrpSerAla 575
 Db 2885 AGAATGTCCTCATGTGTGACCCCAAAATTCAGTAAATTTTAAATCACTACACTAGCGT 2944
 QY 576 ProLeu11eG1yVal1y1Trh1eG1y1yThr-----GlnVal1ProG1n1PheLeu11eSer 592
 Db 2945 CCATCCCAAGAGAAGAAACGATACCTGTGAATGGAGATCAGACAGGACTCATTAAG 3004
 QY 593 Thr-----SerAsn1y1TrhLeu1y1e1yPheSer1TrhAsp1y1ySer11eSer 608
 Db 3005 ACATTGAAACATGACCAATGCTGTGAAGACCTTGAATTAAGAGCCCATGTATTCT 3064
 QY 609 Asp11eG1yPheG1n1e1yA1yG1y1Trh11eThrLeuG1n1SerAspHisCys----- 626
 Db 3065 -----TTCACACTGCTCGGAAACAGTGTGTGCTGACAGCAATTCCTCGAACA 3115
 QY 627 -----LeuAspProG1y11eProVal1AsnG1yG1nA1y----- 637
 Db 3116 GAAAAGCTTTTCTTCTGTGACAGCAGGCTGTGCTGAGAGGGGCGCATGTGTCAAC 3175
 QY 638 -----HisG1yAsnAspPhe1y1yA1yG1yAla1yVal11TrhPheSerCysAspSerG1y 655
 Db 3176 TGCCCCCTGGGAACCTTACTCTTGTGAGGATTCACCTGTGAAGTGTGCTCATGGA 3235
 QY 656 Tyr1TrhLeuSerAspG1yG1uProLeuG1u1yCysG1u-----Pro 668
 Db 3236 TCTTCAACAAAGTGAAGAGGAGCTGGAATGCAAGCTCTGTCCCAAGACTACAGCG 3295
 QY 669 AsnPheG1n1TrpSerA1yA1yLeuProSerCysG1uA1yA1yCysG1yG1yPhe11eG1n 688
 Db 3296 GAATACCTCCATTCAGAAAGGCTCTGAATGCAAAAGCTCAGTGAACAGACGACTAC 3355
 QY 689 G1ySerSerG1y1y----- 692
 Db 3356 TCTTCAAGTGGCTGAGAGACTGGAATGTGTGCTGGTGTACTTATCAACCGAATTT 3415
 QY 693 -----Thr11e----- 694

```

Db      3416 GGATCCGGAGCTGCTCTCTATGCCAGAAACACACACGCTGAAAAGAGACCGT 3475
Qy      695 -----Leuser 696
Db      3476 GACATCTCTGCTTGAGAGTCCCTCCAGTAGAGAAATTCCTCCGTTCTGGCTAA 3535
Qy      697 ProlylPhePro-----AspPheTyrProAsnPro-----Leu 707
Db      3536 CCCGCTACCTCTGCTCCCTCGAGACTATTACCAACCCATGCAAGGAACTCTCTGCTC 3595
Qy      708 AsnSerThrTyr----- 711
Db      3596 GCTTGCTCTTTATGAACTACCAACATCACTGCGCCAGCTTCATCAGACATGCTCA 3655
Qy      712 -----IleIleGluThrSer 716
Db      3656 AGTTTAGCTCTACTTCTCAGACAGAAAGAAAGATAGTCCCTCTGCGCCCTGGA 3715
Qy      717 HisGlySerGlyValPhe-----PheThrPheHisThrPheHisLeu----- 730
Db      3716 CATTCGCCAACAAGATGCAAGTAGCAGCAGTCTTTTACGAAATGCTTTTAAACCCC 3775
Qy      731 -----GluSerGlyHisAspTyrLeuLeuIleThrGluAsn 742
Db      3776 TGCACAAACAGTGAACCTGCCAACAGCTTGGCGGTGTATGTCTGTCTGCCACCT 3835
Qy      743 GlySerPheThrGlnProLeuArgGlnLeuThrGly----- 754
Db      3836 GGA-----TACACAGGCTTAAAGTGTGAACAGATATTGAT 3871
Qy      755 -----SerArgLeuProAlaProIleSerAlaGlyLeu-----TyrGlyAsn 768
Db      3872 GAATGCAAGCTCTGCTTGC-----CTCATGTGTGAATTTGTAGAACCAAGTTGGGGA 3928
Qy      769 PheThrAlaGlnValArgPhe----- 775
Db      3929 TTTCACGTGCGAATGTCTATTGGGCTATTACAGTCAAAATGTGAAGAAATATTAATG 3988
Qy      776 -----IleSerAspPheSerMetSerTyr----- 783
Db      3989 TGTATTCACAGCCCTTGCTTAATAAAGAACTGCACTGACGCGCTTGCAACGTAACGC 4048
Qy      784 -----GluGlyPheAsnIleThrPheSerGlyTyrAspLeuGlnProCysGly 799
Db      4049 TGTACTGTGTGAAGATACATGAGTGTGCACTGTGAACAGCTGAATGAATGCCAG 4108
Qy      800 GluProGluValProAlaTyrSerIleArgGlySerGlnPheGlyValGlyAspThr 819
Db      4109 TCAAGCCCTGCTTAACACACGCACTTTGTAAAGCAAGTTGGGGG----- 4156
Qy      820 LeuThrPheSerCys-----PheProGlyTyrArgLeuGlnGlyThrAla 834
Db      4157 -----TTCTGTGTCAAATGCCCAACCGGATTTTGGGTACTCGGTGTGAAAAAATGTG 4210
Qy      835 -----ArgIleThrCysLeuGlyValArgArgArg 844
Db      4211 GATGAGTGTCTCACTACGACCATGCCAAATGAGAGCACTTGTAAAGATGTGCCAACGC 4270
Qy      845 LeuThrPheSerProLeuProArgCysValAlaGluCysGlyAsnSerValThrGlyThr 864
Db      4271 TTC-----AGGTGT-----CAATGTCCAGCAGGCTTACACGGGACA 4306
Qy      865 GlnGlyThrLeuLeuSerProAsnPheProValAsnTyrAsn----- 878
Db      4307 CACTGTGAACG-----AACATCAACGAGTGCAGTCCAAACCG 4345
Qy      879 -----AsnAsnHisGlyCysIle-----TyrSerIleGlnThrGlnPro 891
Db      4346 TGTAGAAACAGGCGCACTGTGTGATGAACTAAATCATACAGTGTGAATGTCAAGCA 4405
Qy      892 GlyTyrSerGlyIleGlnLeuValAlaArgAlaPheGluLeuSerGlnGlyAspValLeuLys 911

```

```

Db      4406 GGA-----TTTTCAGGCCACAGGTGTGAGACAGAACAGCT----- 4441
Qy      912 ValTyrAspGlyAsnAsnAsnSerAlaArgLeuLeuGlyValPheSerHisSerGluMet 931
Db      4442 -----TCCGATTTTAACTCGGATTTTGAAGTTTCTGGCATCTACGGGTAGCTCTGTA 4495
Qy      932 MetGlyVal-----ThrLeuAsnSerThrSerSerLeuThrPleu-----Asp 946
Db      4496 GATGAATGCTGCCAACCTCTCATGCCGTAACTCGGCATTTCTGGATGAATCTCTGAT 4555
Qy      947 PheIleThr-----AspAlaGluAsnThr----- 954
Db      4556 GTATCAACTACGGAGGCCCATCTCCATGCACTTAGAGTAGCAAAAGAACACCTTC 4615
Qy      955 -----SerLysGlyPheGluLeuHisPheSerSerPheGluLeuIle----- 968
Db      4616 CTCCTGACATGATCAACAGCGTGGTCTTTATGTGAATGAAAGAAAGATCAACCAAC 4675
Qy      968 ----- 968
Db      4676 TGCCCTCCGTAATGATGGCATTGGCATCATTTTGCAATCACAATGACAATTTGGT 4735
Qy      969 -----LysCysGluAspProGlyThr----- 975
Db      4736 GGAACCTGAGAGGCTATATATAGATGGGAATTAATCTACGCGTGTATCGGCTCCATT 4795
Qy      975 ----- 975
Db      4796 GGCAAAAGCATACCTGTGGCGGTGATAGTTCTTGGGCAAGAGCAAAAGAAAGGA 4855
Qy      975 ----- 975
Db      4856 GAGGGTTCAACCCGCTGAGTCTTTTGGGCTTCATTAAGCAAGCTCACTTGAGAC 4915
Qy      975 ----- 975
Db      4916 TATGCTGTCTCCACAGCAGTGAATGTGTGGCCAAGTCTCCGCAAGGAATGAGT 4975
Qy      976 -----ProLysPhe-----GlyTyrLysVal 982
Db      4976 CGGGAAACGTGTAGCATGCGCCGATTTCTGTGGGAATCACGGGGAAAGTAGAGTT 5035
Qy      983 HisAspGlnGlyHisPheAla----- 989
Db      5036 GATTCCAGCAGCATGTTCTGCTGTATTGTCCTGTTTGAAGAGATCCGTCCACCTG 5095
Qy      990 -----GlySerSerValSerPheSerCysAspProGly 1000
Db      5096 AGACCTGCATCAGAAATCGAAAGCCAGGCTCCAAAGTCAAGTCTGTGATCCGGGC 5155
Qy      1001 TyrSerLeuArgGlySerGlnGluLeuLeuCysLeuSerGlyValArgArgThrTyrAsp 1020
Db      5156 TTCCAGATGTTGGGAATCTCTGTGACATATGTGTGAAC-----CAAGGCACTGACCA 5209
Qy      1021 ArgProLeuProThrCysVal-----AlaGluCysGly-----GlyThrValArgGly 1036
Db      5210 CAACCATCTCCCCCACTGTGAACGATTCGCTGTGGCTCCGCTCCGCTTGAGAAATGGC 5269
Qy      1037 GluValSerGlyGlnValLeuSerProGlyTyrProAlaProTyrGlnHisAsnLeuAsn 1056
Db      5270 TTCTACTCAGCGAGCACTTCATGCGGGGACGACGAGTACTGAG----- 5317
Qy      1057 CysIleThrThrIleGluAlaGluAlaGlyCysThrIleGlyLeuHisPheLeuValPhe 1076
Db      5318 -----TGCAACCAAGTGC-----TACTACTCTGCTGGGT 5344
Qy      1077 AspThrGlnGluValHisAspValLeuArgIleThrAspGlyProValGluSerGlyVal 1096
Db      5345 GATTCCGGAATGTTCTGCRACAGACACGAGAGCTGGAACGCAATTTCAACATCTGTCTC 5404
Qy      1097 LeuLeuLysGluLeuSer----- 1102
Db      5405 GATGTGCATGATGTGTGACGTGCGCTCGGACTGTAGTACGACGCTCTGCTGAACACC 5464

```


QY 1102 ----- 1102
 Db 5465 AACGATCTACGATGCTCTGTACACCCACCATACAGGAGATGGAAAACTGTGCA 5524
 QY 1103 ----- GlyProAlaLeuProLysAspLeuHisSerThrPheAsnSer 1116
 Db 5525 GAACTGTAAATGATGAGCTCCAGAAAATCCAGAAAATGGCCCTCT 5572
 QY 1117 ValValLeuGlnPheSerThrAspPheThrSerLysGlnGlyPheAlaIleGlnPhe 1136
 Db 5573 ----- TCTGGCAGATTTACACC-----GTGGTACTGCAGTCACATTT 5611
 QY 1137 Ser-----ValSerThrAlaThr----- 1142
 Db 5612 TCCGTGTACGAAGGGACAGACGTGTGGAGATGACACATCACGTGTTTGGACACTGCC 5671
 QY 1143 -----SerCysAsnAspProGlyIlePro 1150
 Db 5672 GAGTGGATCCCTCAGGCCCTCTGTGAAGCCATTTCTGTGGTGTGCCACCTGTTCT 5731
 QY 1151 GlnAsnGlySerArgSerGlyAspSerTrpGlnAlaGlyAspSerThrValPheGlnCys 1170
 Db 5732 GAAATAGTGTTGTTGACGGGCTCGGCATTCACATATGACAGTAAGTGTGTACAGGTGT 5791
 QY 1171 AspProGlyTyrAlaLeuGlnGlySerAlaGlnIleSerCysValLysIleGlnAsnArg 1190
 Db 5792 GATAAAGATATCTTGTGTGGGATGAAAGTACACATGCTCTGTACTAGTGTCC-- 5848
 QY 1191 PheSerThrGlnProSerProProThrCys--IleAlaProCysGlyGlyAspLeuThr 1209
 Db 5849 -----TGGAGTCATTTCTCTCTGTGTGGGCTAGTGAAGTGT----- 5887
 QY 1210 GlyProSerGlyValIleLeuSerProAsnTyrProGlnProIleProGlyLysGln 1229
 Db 5888 -----TCCAGCCTTAGACACATAAATACGGCAAA-- 5917
 QY 1230 CysAspTrpLysValThrValSerProAspTyrValIleAla-----LeuVal 1245
 Db 5918 -----TACATCTTAAGTGGGCTCACCTACCTTTCT 5947
 QY 1246 PheAsnIlePheAsnLeuGlnProGlyTyrAspPheLeuHisIleTyrAspGlyArgAsp 1265
 Db 5948 ATTGCACTGATCTCTGTGAGAACGATACAGTTTA----- 5983
 QY 1266 SerLeuSerProLeuIleGlySerPheTyrGlySerGlnLeuProGlyArgIleGlnSer 1285
 Db 5984 -----CAGGGCCCATCCCTCTTGAATGC 6007
 QY 1286 SerSerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPhe 1305
 Db 6008 ACAGCTTCCGCGCAGCTGG-----GACAGAGCCGACCTAGCTGCAACTT 6052
 QY 1306 ValIleAspTyrThrGlnAsnProArgGlnSerCysPheAspProGlySerIleLysAsn 1325
 Db 6053 GTC-----TCTGTGGAGAGCGCTCCCAATGCAAAAGAT 6085
 QY 1326 GlyThrArgValAlaSerAspLeuLysLeuGlySerSerValThrTyrCysHisGly 1345
 Db 6086 GCTGTCATCACTGGAGCAACTTCACTTTGGGACACAGTGTCTTAACATGCAAAAG 6145
 QY 1346 GlyTyrGlnValGlnGlyThrSerThrLeuSerCysIleLeuGlyProAspGlyLysPro 1365
 Db 6146 GGCTACACCTTCTGCGGCTGACACCATCATATGC-----CAGGCAACGGCAAA-- 6196
 QY 1366 ValTrpAsnAsnProArgProValCysThrAlaProCysGlyGlyGlnTyrValGlySer 1385
 Db 6197 -----TGAATTCAGTAACACACAGTCTGGCT----- 6226
 QY 1386 AspGlyValValLeuSerProAsnTyrProGlnAsn-----TyrThrSerGlyGlnIle 1403
 Db 6227 -----GTCTCTGTGTACGAGCCCCCAATGTGACACAGCCCTCTCCAGAGACT 6274

QY 1404 CysLeuTyrPheValThrValProLysAspTyrValValPheGlyGlnPheAlaPhePhe 1423
 Db 6275 GCT-----CAGAGCTCTTTGGAGACACCGGTTTTC 6307
 QY 1424 HisThrAlaLeuAsnAspValValGluValHisAspGlyHisSerGlnHisSerArgLeu 1443
 Db 6308 TACTGTGCG-----GATGGCTACAGC----- 6328
 QY 1444 LeuSerSerLeuSerGlySerHisThrGlyGlnSerLeuProLeuAlaThrSerAsnGln 1463
 Db 6329 -----CTGGCTGATATATTCACG 6346
 QY 1464 ValLeuIleLysPheSerAlaLysGlyLeuAlaProAlaArgGlyPheHisPheValTyr 1483
 Db 6347 CTCATGTCAATGCCAGGGGAACTGGGTTCCCGCGGGC----- 6388
 QY 1484 GlnAlaValProArgThrSerAlaThrGlnCysSerSerValProGlnProArgTyrGly 1503
 Db 6389 CAGGCTGTCCGGCTGCATAGCTCACTTCTGTGAAAAACCCCAATCTGTTCTTCAAGC 6448
 QY 1504 -----LysArgLeuGlySerAspPheSerValGlyAlaIleValArgPheGlnCys 1520
 Db 6449 ACTTGGAACTGTGTAGCAAAAGATTGACAGCTCGGTGTGTAGCTTCAAGTGC 6508
 QY 1521 AsnSerGlyTyrAlaLeuGlnGlySerProGlnIleGlnCysLeuProValProGlyAla 1540
 Db 6509 ATGAGAGTTTGTCTGTGAACCTCAGGAAGATTGAATGCTTGA-- 6556
 QY 1541 LeuAlaGlnTrpAsnValSerAlaProThrCysValValProCysGlyGlyAsnLeuThr 1560
 Db 6557 -----GGTGAAGATGGAGC 6571
 QY 1561 GluArgArgGlyThrIleLeuSerProGlyPheProGlnProTyrLeuAsnSerLeuAsn 1580
 Db 6572 -----CTTCTCCCTC-----TCGGTCCAG 6592
 QY 1581 CysValTrpLysIleValValProGlnGlyAlaGlyIleGlnIleGlnValValSerPhe 1600
 Db 6593 TGC----- 6595
 QY 1601 ValThrGlnGlnAsnTrpAspSerLeuGlnValPheAspGlyAlaAspAsnThrValThr 1620
 Db 6595 ----- 6595
 QY 1621 MetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerThrSerAsnGln 1640
 Db 6596 -----ATCCGAGTG----- 6604
 QY 1641 LeuTyrLeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHisIleuGlnTyr 1660
 Db 6604 ----- 6604
 QY 1661 LysThrValGlyLeuSerSerCysProGlnProAlaValProSerAsnGlyValLysThr 1680
 Db 6605 -----CGATGGGAGAGGCTCCACATCGCAATGGTCAACCGAGT 6646
 QY 1681 GlyGluArgTyrLeuValAsnAspValValSerPheGlnCysGlnProGlyTyrValLeu 1700
 Db 6647 GGGAAACATACAGTTTGGGCGCGTGTGGCTTCAAGCTCCACAGAGATTTCTATATC 6706
 QY 1701 GlnGlyHisValHisIleSerCysMetProGlyThrValArgArgTrpAsnTyrProPro 1720
 Db 6707 AAGGGGAGAGAGAGACAGCTGT-----CAGGCAACAGACGTGAGTAACCCACG 6760
 QY 1721 ProLeuCysIleAlaGlnCysGlyGlyThrValGlnGlnMetGlnGlyValIleLeuSer 1740
 Db 6761 CCCACTGC----- 6769
 QY 1741 ProGlyPheProGlyAsnTyrProSerAsnMetAspCysSerTrpLysIleAlaLeuPro 1760
 Db 6770 -----CATCT-----GTGCTGTAAACAGACCACTTAAGTTAG 6805
 QY 1761 ValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGlnProAsnHisAspTyr 1780

D	6806				6865	D	7589	-----	GTGGGCAATGCTACCACTCTGTGGG	7615	
Q	1781				1800	Q	2134	ProSerGlyGln-----	SerProLeuLeuValAlaLeuSerGlyAsnTyrSer	2149	
D	6866	-----	-----	-----	6892	D	7616	GAATAATGGCCAGTGGCTCGAGAGAAACCAATGTGCMAACCCATTGAA	-----	7663	
Q	1801				1820	Q	2150	AlaProLeuIleValThiSerSerSerAsnSerValTyrLeuArgTyrPserSerAspHis	2169		
D	6893	-----	-----	-----	6934	D	7663	-----	-----	7663	
Q	1821				1840	Q	2170	AlaTyrAsnArgLysGlyPheLysIleArgTyrSerAlaProTyrCysSerLeuProArg	2189		
D	6935	-----	-----	-----	6964	D	7664	-----	-----	7678	
Q	1841				1860	Q	2190	AlaProLeuHisGlyPheIleLeuGlyGlnThiSerThiGlnProGlySerIleHis	2209		
D	6965	-----	-----	-----	7024	D	7679	GAGATTATTAATGCG-----	CAATTCTCTCCGTGAGCTTTCAGATAGACAAACCATTCACA	7735	
Q	1861				1880	Q	2210	PheGlyCysAsnAlaGlyTyrArgLeuValGlyHisSerMetAlaIleCysTyrArgHis	2229		
D	7025	-----	-----	-----	7084	D	7736	TACTTTGTGACCGGGCTTCGGCTCGAAGTCCCAATCCCTGACCTTTTAAGACA	7795		
Q	1881				1898	Q	2230	ProGlnGlyTyrHisLeuTyrPserSerGlyAlaIleProLeuCysGlnAlaLeuSerCysGly	2249		
D	7085	-----	-----	-----	7135	D	7796	GGTGCAC-----	-----	7843	
Q	1899	-----	-----	-----	1917	Q	2250	LeuProGlnAlaProLysAsnGlyMetValPheGlyLysGlyTyrThrValGlyThrLys	2269		
D	7136	-----	-----	-----	7183	D	7844	GACCCACAGCCCATGTAATAATGTTTCGTGAAGGTCCGATTAACAGTACGGGCCATG	7903		
Q	1918	ProTyrSerSerSerGlnAspCysValTyrLeuIleThrValProIleGlyHisGlyVal	1937	D	7237	D	2270	AlaValTyrSerCysSerGlnGlyTyrHisLeuGlnAlaGlyValAlaThrAlaGlu	2289		
D	7184	-----	-----	-----	7237	D	7904	ATCATTAATAGCTGCTTCCCTGGTTTCAGGTCTGGTATGATGCACGACC-----	7957		
Q	1938	ArgLeuAsnLeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTyrAsp	1957	D	7282	Q	2290	CysLeuAspThrGlyLeuTyrPserAsnArgAsnValProProGlnCysValProValThr	2309		
D	7238	CAAA-----	GGCCCTCTGTCTCGAAGTGCCTGCCATCCGGCAA-----	7282	D	7958	TGTGAAGATCGGGA-----	TGTCACAGCTCAGC-----	CCAACTGTGATCCCATTAAC	8008	
Q	1958	GlyProGlnGlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaLysLysThr	1977	D	7285	D	2310	Cys-----	ProAspVal-----	SerSerIleSerValGluHisGlyArgTyr	2323
D	7283	GGT-----	-----	7285	D	8009	TGGCGTCTCCCTCCATACATGAGCTTGTGACTGATTAAGTACAGAGATGGCCAGGA	8068			
Q	1978	ValGlnSerSerSerAsnGlnValLeuLeuLysPheHisArgAspAlaAlaThrGlyGly	1997	D	7285	D	2324	ArgLeuIlePheGlu-----	-----	-----	2328
D	7285	-----	-----	7285	D	8069	CATTGTGATCAAGAAATGACATGATGGAATGCCATATGTGGCTACCTCAACATTTG	8128			
Q	1998	IlePheAlaIleAlaPheSerAlaTyrProLeuThrLys-----	CysProProPro	2014	Q	2329	-----	-----	ThrGln	2330	
D	7286	-----	-----	7286	D	8129	GAAACAACAGCTAAGGCTTGGAAATACAAAGAGTGGCTGCTCATGATGCCAC	8188			
Q	2015	ThrIleLeuProAsnAlaGluValAlaThrGluAsnGluGluPheAsnIleGlyAspIle	2034	D	2331	TyrGlnPheGlnAlaGlnLeuMetLeuIleCysAspProGlyTyrTyrTyrThrGlyGln	2350				
D	7325	CCCTGATCCCTCGGC-----	GTCCTGCGCTTCGCGGTGCTTCAATTTGGCAGTACT	7381	D	8189	TTCCTATGAGCAGATGTTCTCTACAGCTGGAGCTGTGATGAATGCTGGGATTC	8248			
Q	2035	ValArgTyrArgCysLeuProGlyPheThrLeuValGlyAsnGluIleLeuThrCysLys	2054	D	2351	ArgValIleLeuArgCysGlnAlaAsnGlyLysTyrPserIleuLysPserThrProLys	2370				
D	7382	GTCMAAGATCTGTGTGACGCGGTTTCTTAAAGACAGTCCAACATCTCTGCGAG	7441	D	8249	CTGTGCTGATCTGCACGAGAAATGACATGAAAT-----	GGTACCGCACTCTTGC	8302			
Q	2055	LeuGlyThrTyrLeuGlnPheGluGlyProProProLys-----	GluValHisCysPro	2073	D	2371	ArgIleIleSerCysGlyGlnLeuProIle-----	ProProAsnGlyHisArgIleGlyThr	2389		
D	7442	GCTGATATGC-----	ACCTGAGATTTCTTCATGCCCCGAATGCCGTTTCCGGTGAATGATCCC	7495	D	8303	ATTTCATTGAATGT-----	GATTTGCCGTGCTCCGAAATATGCTTTTAACATTTGACA	8359		
Q	2074	ThrAsnGluLeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTyrProGlySer	2093	D	2390	LeuSerValTyrGlyAlaThrAlaIlePheSerCysAsnSerGlyTyrThrLeuValGly	2409				
D	7496	CAACTGAGAGATCTCAAC-----	GGTATCATCCACGTAACAGGCTTGCTGCTTACTTC	7549	D	8360	CAGACGACTATGGGACAGTCTGCACAAATATAGCTGCAGCGGGGCACATTTCAAGAGCG	8419			
Q	2094	TyrProGlnPheGlnThrCysSerTyrPheValArgValGluProAspTyrAsnIleSer	2113	D	2410	SerArgValArgLysCysMetAlaAsnGlyLeuTyrPserGlySerGluValArgCysLeu	2429				
D	7550	AGCACCAAGCTTACACCTGC-----	-----	7588	D	8420	TCCACTTAAGACTCTGTGTGCAGAAATACAGAGGTGGCACTGTTTCAAGCTGTGAA	8479			
Q	2114	LeuThrValGluTyrPheLeuSerGlyLysGlnTyrAspGluPheGluIlePheAspGly	2133	D	2430	AlaGlyHisCysGlyThrProGluProIleValAlaAsnGlyHisIleAsnGlyLysAsnTyr	2449				
					D	8480	GCCATCTCATGCAATAGCAACCAACCACTGGATGATGATCCATCAAGAGATGATAC	8539			

QY 2450 SerTyrArgGlySerValValTyrGlnCysAsnAlaGlyPheArgLeuIleGlyMetSer 2469
 Db 8540 TCCTACCTGGGTGGTGTATCTACTACGAGTGTGACTCTGGCTATATCTCAATGGCTCTAAG 8599
 QY 2470 ValArgTleCysGlnGlnAspHisHisTyrSerGlyLysThrProPheCysValProIle 2489
 Db 8600 AAGAGCATGTCCCAAGAAATAGATGGATGGGATGGAGCCCATGTGATATCTCGTA 8659
 QY 2490 ThrCysGlyHisProGlyAsnProValAsnGlyLeuThrGlnGlnAsnGlnPheAsnLeu 2509
 Db 8660 GACTGTGCTCACCCTCCAGTCCCAAGATGGCCGAGTGAAGGAGAAATACATCCATTC 8719
 QY 2510 AsnAspValValLysPheValCysAsnProGlyTyrMetAlaGlnGlyAlaAlaArgSer 2529
 Db 8720 CAAAAGAGATATACATCTCTGCTCCGTAAGGGTTCATCTGAAAGAGCCAGAGTCTGT 8779
 QY 2530 GlnCysLeuAlaSerGlyGlnTyrPheSerAspMetLeuProThrCysArgIleIleAsnCys 2549
 Db 8780 ATCTGTCTTACCAATGGAGATGGAGTGGTCCACTCCAGCTGCATGCTGTATGATGT 8839
 QY 2550 ThrAspProGlyHisGlnGlnLysSerValArgGlnValHisAlaSerGlyProHisArg 2569
 Db 8840 CTGCCCCACACAGTGCCTCAATGGGGTG-----GCAGATGGCTTAGAC-- 8884
 QY 2570 PheSerPheGlyThrThrValSerTyrArgCysAsnHisGlyPheTyrLeuLeuGlyThr 2589
 Db 8885 TATGGGTTCAAGAAAGAGATAGCGTTCCACTGTCTAAGGGCTATGTCTGCAGGGGCT 8944
 QY 2590 ProValLeuSerCysGlnGlnLysArgGlyThrTyrAspArgProArgProGlnCysLeu 2609
 Db 8945 CCAGAGCTTCACTGTCTCAATGGAGATGGAGCTTGGGATCCAGAGTCTCTGTAAACA 9004
 QY 2610 ValSerCysGlyHisPheProGlySerProProHisSerGlnMetSerGlyAspSerTyrThr 2629
 Db 9005 GCTACCTGTGTCTCTCTGCCGACCTTCCCAAGGGCTCCCTATATGGCTTTCTTTTAT 9064
 QY 2630 ValGlyAlaValAlaArgTyrSerCysIleGlyLysArgThrLeuValGlyAsnSerThr 2649
 Db 9065 CATGGGGGCCACATACAGTATCACTGTCTTACGTATTAAGTTCAATGGAATCCATCA 9124
 QY 2650 ArgMetCysGlyLeuAspGlyHisIleThrThrGlySerLeuProHisCys----- 2665
 Db 9125 AGAAGATGCTTCCCATGTGTCTGTGAGGCGACGCTCGCCATCTGCTTACCTTCAG 9184
 QY 2666 -----SerGlyThrSerValGlyVal----- 2672
 Db 9185 TGTTCCACACCCCATCATTTCAACAGGAAACCATCAACCCACATGATTTGGATGTGAAAG 9244
 QY 2672 ----- 2672
 Db 9245 ACGGTCCAGATGATGCTTCAAAAGGCTTCAGGTGCTTGAATTCACACTGT 9304
 QY 2673 -----CysGlyAsp----- 2675
 Db 9305 GATGCCAATGGCAATGTGTGACGTCCACTGTGTAGACACGCTCAGTGGGGGCTCTC 9364
 QY 2676 ProGlyIleProAlaHisGlyIleArgLeuGlyAspSerPheAspProGlyThrValMet 2695
 Db 9365 CCAACCATACCC---AAGCAATGTGTCTTGAAGGAGCGCTTTCGAGAGCAATGTGTGA 9421
 QY 2696 ArgPheSerCysGlnLysGlyHisValLeuArgGlySerSerGlnArgThrCysGlnAla 2715
 Db 9422 ACTTACAGCTGCAGACCTGGCTACACCATCAAGAGTATGTTCAATCTGATTTGTACGAA 9481
 QY 2716 AsnGlySerThrPheSerGlySerGlnProGlyValIleSerCysGlyAsnProGly 2735
 Db 9482 AAAAGCATATGGAGCCAGCTTACCAACGCTGTGAACCCCTGTCTGTGACCCCAACA 9541
 QY 2736 ThrProSerAsnAla----- 2740
 Db 9542 ACTGTAGCCATCATCGTGGCAACAGAGAGGCTCATCTATGAAAGCAAAGTGAATCTC 9601

QY 2740 ----- 2740
 Db 9602 AGGTGTCTGAAAGGATATGTGATTTGGATTACAGATATCATTTACCTGCCAGCAAGAT 9661
 QY 2740 ----- 2740
 Db 9662 GGCATTGGGTCCCTGAAGATATCACCTGACGTCTTAAAAATATGCCCTGTGCATCCAA 9721
 QY 2741 -----ArgValAlaPhe-----SerArgGlyLeuValPheSerSerSerIleValTyrGlu 2757
 Db 9722 ATGACACGATATCGTTTACGAGATGACTTCCAGGTGAACAGACATTTCTGTGTCA 9781
 QY 2758 CysArgGlnGlyTyrTyrAlaThrGlyLeuLeuSerArgHisCysSerValAsnGlyThr 2777
 Db 9782 TGTGAGAAAGGTTTATCCACAGAGAGTGAATGTGTCAACATGCCCAGCCGATACA 9841
 QY 2778 TTPThrGly-----SerAspProGlyCysLeuValIleAsnCysGlyAspProGlyIle 2795
 Db 9842 TGGAGCCACACATTTTCTGATGAATCTGTATCCAGTGTGTTGTGGGATCTGAAAGC 9901
 QY 2796 ProAlaAsnGlyLeuArgLeuGlyAsnAspPheArgTyrAsnLysThrValThrTyrGln 2815
 Db 9902 CCAAGCATGGTCTCGGTGGTGGCAATAGACACAGCTTTGGAAGCACCATTTTATCCAG 9961
 QY 2816 CysValProGlyTyrMetMetCysLeuSerHisArgValSerValLeuSerCysThrLysAsp 2835
 Db 9962 TGTGACCTGTGCTACAAATTAAGAGGGAACAGGAAACGAATC-----TCCAGAGAAAC 10015
 QY 2836 ArgThrTyrAsnGlyThrLysProValCysLysAlaLeuMetCysLysProProProLeu 2855
 Db 10016 AACACGTGAGAGTGGAGAGGTGACAGTGTGCAGAGAAACAGATGTGACACTCAGCTGAG 10075
 QY 2856 IleProAsnGlyLysValValGlySerAspPheMetTyrGlySerSerValThrTyrAla 2875
 Db 10076 TTTCCAAATGGGAAGCTGTCTTGAAGAACACCAATCTGACCCAGCTTGTGTTTCT 10135
 QY 2876 CysLeuGlnGlyTyrGlnLeuSerLeuProAlaValPheThrCysGlnGlyLysAsnGlySer 2895
 Db 10136 TGTCAACAGAGCTTACCACTGGAAGGTGTCCTCCAGGACACATGCACTGCATCCAAATGGAAC 10195
 QY 2896 TTPThrGlyGlnLeuProGlnCysPheProValPheCysGlyAspProGlyVal---Pro 2914
 Db 10196 TCGAATCACTGACCTCCCTCGCAACCAATCATGCTCCGTCTTGTGTGATTCCT 10255
 QY 2915 SerArgGlyArgArgGlnAspArgGlyPheSerTyrArgSerSerValSerPheSerCys 2934
 Db 10256 GAGAAAGCGGCTCTTCTGAAAGAGATTTATGTGACCCAGAAATGATCTATCAAGTGC 10315
 QY 2935 HisProProLeuValLeuValGlySerProArgArgPheCysGlnSerAspGlyThrTyr 2954
 Db 10316 AAGGAAGGCTTCTGCTCAAAAGGCAATGTGTATCATCAGTGAAGCCGTGAGAGACATGG 10375
 QY 2955 SerGlyThrGlnProSerCysIleAspProThrIleThrCysAlaAspProGlyVal 2974
 Db 10376 ACGCACCAATGCCAGATGT-----GAAAAATCTCTGTGTCTTCAAGTCAAC 10426
 QY 2975 ProGlnPheGlyIleGlnAsnAsnSerGlnGlyTyrGlnValGlySerThrValLeuPhe 2994
 Db 10427 GTAGGAATGCATTTGCTCGAGAGTGTAT---TTCAGATATGGAGAAATGATCACTTAC 10483
 QY 2995 ArgCysGlnLysGlyTyrLeuLeuGlnGlySerThrThrArgThrCysLeuProAsnLeu 3014
 Db 10484 TCCTGTACAGTGGCTACATGCTTAAAGGTTCCCTCCGAGAGTGTGTGCTTGAAGAAATGGA 10543
 QY 3015 ThrTyrPheSerGlyThrProAspCysValProHisHisCysArgGlnProGlyThrPro 3034
 Db 10544 ACATGG-----ACACCATCTCTGTGTGAGAGCTGTGTGTGCTTCCCA----- 10588
 QY 3035 ThrHisAlaAsnValGlyAlaLeuAspLeuProSer----- 3046
 Db 10589 ---TGTCAAAATGAGAGTGTCTGTACACGTCCAAATGCTTGTCTATGCCAGACGGCTGG 10645
 QY 3047 MetGlyTyrThrLeuIleThrProAlaArgArgAlaSerProSerArgValAlaProSer 3066

```

Db      10646 ATGGGACGTCCTCTGT-----GAGAGCCCAATATGCACTACCTCCCTGT 10687
Qy      3067 ThrAlaProAlaArgMetAlaAlaGlyAlaAlaSerArgProSerAlaTyr 3084
Db      10688 TTGAATGATGGGCGCTGTGTGGCCCTTATCATGATGTAAGTACTGCCACAGGCTGG 10741

```

Search completed: October 19, 2004, 00:19:28
 Job time : 3393 secs

This Page Blank (uspto)

92 LeuGlySerSerValGlnPheThrCysAsnGluGlyTyrAspLeuGlnGlySerLysArg 11

```

Db 1370 TTCAATGCCGCTGTGGGGTCCGATGTGCGCCGGGCTTTGACCTTGTGGAGAAGACGATC 1429
QY 112 ILeThrCysMetIysValSerAspMetPheAlaIlePheSerAspHisArgProValCys 131
Db 1430 CATTGTGTCAACC-----AATGTTTGTGTCTGGGACAGAAAGCTTGTGC 1477
QY 132 ArgAlaArgMetCysAspAlaHisLeuArgGlyProSer---GlyIleIleThrSerPro 150
Db 1478 AGAGTGAAGACGTGC-----CCCACTCTCGACAGCCCAACAGCCGACATGAGTGTCC 1534
QY 151 AsnPheProIleGlnIlyAspAsnAsnAlaHisCysValIlePheIleIleThrAlaLeuAsn 170
Db 1535 ACTCGGAATAATGCTCTC-----AACACCTGTGT-----TTGGTTACCTGC----- 1576
QY 171 ProSerIysValIleIysLeuAlaPheGluGluPheAspLeuAlaArgGlyIlyAspThr 190
Db 1577 -----AATGAAGATACAGATT 1594
QY 191 LeuThrValGlyAspGlyGlyGlnAspGlyAspGlnIysThrValLeuIlyMetSerGln 210
Db 1595 -----GAAGCAGCACTAGGCTTACCTGTCAAGA 1624
QY 211 AsnAlaCysSerAspSerProHisThrProGlySerArgIleProGlyIleMetSerGly 230
Db 1625 AATGCCAGTGGAGTGGCCAG-----CCCGGTGTGTGA----- 1663
QY 231 AspIleThrArgGlnIlyStrpThrValLeuGluIleCysArgAspIleSerSerSerAsp 250
Db 1663 ----- 1663
QY 251 AlaArgSerGlySerValArgIysSerProIlyThrSerAsnAlaValGluLeuValAla 270
Db 1664 ---CGCATTTGTGCCACTTCCAGAACCCCAA----- 1693
QY 271 ProGlyIleThrGluIleGluGlnIlySerCysGly---AspProGlyIleProAlaIlyGly 289
Db 1694 ---GGCTCATCATTTCTCCACCCAGCTGGCGCAAGCGCCGACAGCT----- 1741
QY 290 ArgArgGluGlySerArgPheHisIleGlyAspThrLeuIysPheGluCysGlnProAla 309
Db 1742 -----GGATAGACTGTCAAGCTAGCTGCCCCAGGGA 1774
QY 310 PheGluLeuValGlyGlnIlyValIleThrCysGlnIysAsnAsnGlnIlePheSerArgIly 329
Db 1775 TACATTTTATCCGGGGTCAAGAGAAGTGAATGTGCCACATCTGGAAAGTGAAGTCCCAA 1834
QY 330 --LysProGlyCysValPheSerCysPhe----- 338
Db 1835 GTTCAGACAGCTGTGTCAAAAGATGTGAGGCTCCCAAAATCAGCTGTCCAAATGACATT 1894
QY 339 -----PheAsnPheThrSerProSer--- 345
Db 1895 GAGGCAAAAGACTGGGAGCAGACGACTGTGTAATGCCACTGCGCAAGTCCCAACAGCT 1954
QY 346 -----GlyValValLeuSerProAsnIlyProGlyAspIlyr--- 357
Db 1955 AAAGACAACCTGTGTGAAAAGTGTCT-----AGTCACCTGCCACCCAGCTTTTACCCA 2008
QY 358 -----GlyAsnHisLeuHisCysValIlePheIleIleLeuAla 369
Db 2009 CCTTACCTCTTCCCAATTTGAGACGTGTGCCATCACTACCGCAACCGAATCTATC----- 2064
QY 369 IahArgProGluSerArgIleHisLeuAla-PheAsnAspIleAspValGluPro----- 386
Db 2065 ---CGGTAAACCAAGCAGCTGCACTTCTATTAAGTCAATTGATGTGAACCGCTGTC 2122
QY 387 -----Gln 387
Db 2123 ATAGATTGTGCGGATCTCCACTCCAAATCCAGGTGTAGAGAAGAGACCCCTGCAGAC 2182
QY 388 PheAspPheLeuValIleIysAspGlyAlaIleThrAlaGluAlaIleProValLeuGlyThrPhe 407
Db 2183 TGGGATGAGGCTCAATTTCAACAACTCCGGGGCTGAATTGTGTCAATTACCAAGAGTCAAC 2242
QY 408 Ser---GlyAsnGluLeuProSerSerIleThrSerSerGlyHisValAlaIleGluGlu 426
Db 2243 ACACAAGGGGACATGTTTCTCAATGGGGAACCGGTGTGTGTGTACACAGC----- 2293
QY 427 PheGluIleThrAspHisSerThrGlyIysArgGlyPheAsnIleThrPheThrThrPheArg 446
Db 2294 -----ACTGACCCCTCAGGCAACAAACAGACCTGTGACATC---CACATTGTCAATAA 2344
QY 447 HisAsnGluCysProAspProGlyValProValAsnGlyIysArgPheGlyAspSerLeu 466
Db 2345 GGTTCCTCCGTGTGAGGCTCCCTTCAACCCCTGTAAACGGGGAATTATCTGTGCCAGAT 2404
QY 467 GlnLeuGlySerSerIleSerPheLeuCysAspGluGlyPheLeuGlyIleThrGlnIlySer 486
Db 2405 AGTGTGTGATTACTGTAGCTGTGAGCTGACAGGAGGCTATGATTTCACAGAAAGGTCA 2464
QY 487 ---GluThrIleThrCysValLeuIysGluGly---SerValVal 499
Db 2465 CCGAGAAATGACTACTGTCTTTGAAGATGATGTGAGACACCAATACCTTACAGAA 2524
QY 500 TrpAsnSerAlaValLeuArgCysGluAlaProCysGly----- 512
Db 2525 TGGCCAGACTGTGCTATAAACGTTTGCAAACATGTTCAAGTCTTGAATGCTA 2584
QY 513 -----GlyHisLeuThrSerProSerGly 520
Db 2585 TACAAACCACTCGCTGTGATGACATGATCTGTTAAGAACTTTTGTGACACATTGTAG 2644
QY 521 ThrIleLeuSerProGlyIlyProGlyPheIlyIysAspAla----- 534
Db 2645 ACTACCTGTGGGAACATGTGCTCCGTCTTTGTAAAGATGATGATGACATGACATGAC 2704
QY 535 -----LeuSerCysAlaIleThrValIleGluAlaGlnProGlyIlyPro----- 548
Db 2705 CTGGAGACCTGACCAAAAATATCTGATGATTAATTACACTATGAAATGCTTT 2764
QY 549 -----IleIysIleThrPheAspArgPhe 556
Db 2765 GCAATTGGACCAAGAGGCTGGGGTGCAGGCAACAGGCTGTGATTATCTTACATCACTTC 2824
QY 557 IysThrGluValAsnIlyAspThrLeuGluValArgAspGlyArg----- 571
Db 2825 CTGATGTGTACAGAAACACCAACGATGTGGCAAGCCAGATTCGACAGATTAA 2884
QY 572 ----- 572
Db 2885 AGAAGTGTCCATTGTGTGACCCCAAAATTCAGTAAATTTTAAATCATCAGCTAGCGTG 2944
QY 576 ProLeuIleGlyValIlyThrHisIleGlyThr-----GlnValProGlnPheLeuIleSer 592
Db 2945 CCACCTCCAGAGAAAGAAACATACCTCTTGAATTTGAGAAATCAGACAGGACTCATTAAG 3004
QY 593 Thr-----SerAsnIlyLeuIlyLeuIysPheSerThrAspIysSerHisSer 608
Db 3005 ACATTGAAACATACCAATGCTGCTGAAAGACCTTGAATTAAGAGCCCATGTATTCT 3064
QY 609 AspIleGlyPheGlnLeuArgIlyGluIleThrIleLeuGlnIleSerAspHisCys----- 626
Db 3065 ---TTCACCTGCTCGGAAACAGTGTGAGTGTGACAGCAATTCCTCGAACA 3115
QY 627 -----LeuAspProGlyIlyIleProValAsnGlyGlnArg----- 637
Db 3116 GAAAGGCTTTCTCTTGTGACAGACAGGCTGTGCTGAGAGGGGCGCATGTGTCAAC 3175
QY 638 -----HisGlyAsnAspPheIlyValGlyAlaLeuValIleThrPheSerCysAspSerGly 655
Db 3176 TGGCCCTGGGAACCTTACTCTGTGAGCATTCACCTGTGAAGAGTGTGCTCATGGGA 3235
QY 656 TyrThrIleSerAspGlyGluProLeuGluIysGlu-----Pro 668
Db 3236 TCTTACCAAGATGAAGAAAGGCGACGTGGAATGCAAGCTGTGTCCCCCAAGAGCTCACG 3295

```


Db 5318 -----TGACACGAGTGGC---TACTACTGCTGCGGT 5344
Qy 1077 AspThrGluGluValHisAspValLeuArgIleThrAspGlyProValGluSerGlyVal 1096
Db 5345 GATTCCCGAATGTTCTGCRACGACACGAGGAGCTGGAACGCGATTGCACCATCTGCTC 5404
Qy 1097 LeuLeuGluLeuSer----- 1102
Db 5405 GATTCGATGAGTGTGCGAGTGGCTCGGACTGTAGTACGACGCTCTGCTGAAACC 5464
Qy 1102 ----- 1102
Db 5465 AACGGATCTACGATGCTCTGTAAACCACATACAGGAGATGGAAAACTGTGCA 5524
Qy 1103 -----GlyProAlaLeuProValAspLeuHisSerThrPheAsnSer 1116
Db 5525 GAACCTGTAATGTAAGCTCCAGAAATCCAGAAATGCGCCTCT----- 5572
Qy 1117 ValValLeuGlnPheSerThrAspPheThrSerLysGlnGlyPheAlaIleGlnPhe 1136
Db 5573 -----TCTGGCGAATTACAC-----GTGGTACTGCAATCACAATT 5611
Qy 1137 Ser-----ValSerThrAlaThr----- 1142
Db 5612 TCTGTGACGAGGAGGACGAGCTGTGGAGTAGACCATCAGCTGTTGGAGACTGGC 5671
Qy 1143 -----SerCysAsnAspProGlyIlePro 1150
Db 5672 GAGTGGATGCGCTCAGCGCGTCTGTGAAGCCATTCTGTGGTGTCCACCTGTGCT 5731
Qy 1151 GlnAsnGlySerArgSerGlyAspSerTrpGluAlaGlyAspSerThrValPheGlnCys 1170
Db 5732 GAAATAGTGTGTGACGGGTGCGCATTCACATATGCGAGTAAAGTGTGACAGTGT 5791
Qy 1171 AspProGlyTyrAlaLeuGlnGlySerAlaGluLeuSerCysValLysIleGluAsnArg 1190
Db 5792 GATTAAGGATATCTTGTGTGGGAGTGAAGTCAAGTCACTGTGTAAGTGTCC--- 5848
Qy 1191 PheThrTrpGlnProSerProProThrCys---IleAlaProCysGlyGlyAspLeuThr 1209
Db 5849 -----TGAGTCTTCTCTCTCTGCTGCGGCTAGTGAAGTGT----- 5887
Qy 1210 GlyProSerGlyValIleLeuSerProAsnLysProGluProTyrProGlyLysGlu 1229
Db 5888 -----TCCAGCTGAGGACATTAATATACGGCAAA--- 5917
Qy 1230 CysAspTrpLysValThrValSerProAspTrpValIleAla-----LeuVal 1245
Db 5918 -----TACATCTTAAGTGGGCTCACTACCTACTTCTTCT 5947
Qy 1246 PheAsnIlePheAsnLeuGluProGlyTyrAspPheLeuHisIleTyrAspGlyArgAsp 1265
Db 5948 ATTCATCTGATCTCTCTGGAACGAGTACAGTTTA----- 5983
Qy 1266 SerLeuSerProLeuIleGlySerPheTyrGlySerGlnLeuProGlyArgIleGluSer 1285
Db 5984 -----CAGGGCCCATCCCTCTGGAATGC 6007
Qy 1286 SerSerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPhe 1305
Db 6008 ACAGCTTCCGCGAGCTGG-----GACAGAGCGGACCTGCTCAACT 6052
Qy 1306 ValIleAspTrpThrGluAsnProArgGluSerCysPheAspProGlySerIleLysAsn 1325
Db 6053 GTC-----TCTGTGGAGAGCTCCCAATCGCAAGAT 6085
Qy 1326 GlyThrArgValGlySerAspLeuLysLeuGlySerSerValThrTyrCysHisGly 1345
Db 6086 GCTGTATCATCTGAGGACCACTTCACCTTTGGGAACACAGTCTTACACATCCAAAG 6145
Qy 1346 GlyTyrGlyValGluGlyThrSerThrLeuSerCysIleLeuGlyProAspGlyLysPro 1365
Db 6146 GGCTTACACCTTGTGGGCTGACACCATCATATGC-----CAGGCCAAACGGCAAA--- 6196
Qy 1366 ValTrpAsnAsnProArgProValCysThrAlaProCysGlyGlyGlnTyrValGlySer 1385
Db 6197 -----TGGAAATTCAGTACCAACACAGTGCCTGGCT----- 6226
Qy 1386 AspGlyValValLeuSerProAsnLysProGlnAsn-----TyrThrSerGlyGlnIle 1403
Db 6227 -----GTCTCTGTGACGAGCCCCCAATGTGAGACCAAGCTCTCCAGAGACT 6274
Qy 1404 CysLeuTyrPheValThrValProLysAspTyrValValPheGlyGlnPheAlaPhePhe 1423
Db 6275 GCT-----CACAGGCTCTTTGAGACACCGGCTTTTAC 6307
Qy 1424 HisThrAlaLeuAsnAspValValGluValHisAspGlyHisSerGlnHisSerArgLeu 1443
Db 6308 TACTGTGG-----GATGGCTTACAGC----- 6328
Qy 1444 LeuSerSerLeuSerGlySerHisThrGlyGluSerLeuProLeuAlaThrSerAsnGln 1463
Db 6329 -----CTGGCTATTAATTCACAG 6346
Qy 1464 ValLeuIleLysPheSerAlaLysGlyLeuAlaProAlaArgGlyPheHisPheValTyr 1483
Db 6347 CTCATCTGCAATGCCAGGAGAACTGGGTTCCCGCGGCGC----- 6388
Qy 1484 GlnAlaValProArgThrSerAlaThrGlnCysSerSerValProGluProArgTyrGly 1503
Db 6389 CAGGCTGTCCCGCGCTGACATACCTCTGTGAAMAAACCCCATCTGTTCTTACAC 6448
Qy 1504 -----LysArgLeuGlySerAspPheSerValGlyAlaIleValArgPheGluCys 1520
Db 6449 ATCTTGATCTGTGACGAAACAAAGTTTGCAGCTGCTCGGTAGTACCTTCAAGTGC 6508
Qy 1521 AsnSerGlyTyrAlaLeuGlnGlySerProGluIleGluCysLeuProValProGlyAla 1540
Db 6509 ATGAGAGCTTTGTGTCTGTAACACTCAGCAAGATTGAATGCTGAGA----- 6556
Qy 1541 LeuAlaGlnTrpAsnValSerAlaProThrCysValAlaProCysGlyGlyAsnLeuThr 1560
Db 6557 -----GTGGAGAGTGGAG 6571
Qy 1561 GluArgArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeuAsn 1580
Db 6572 -----CTTCTCCCTC-----TCGGTCCAG 6592
Qy 1581 CysValTrpLysIleValValProGluGlyAlaGlyIleGlnIleGlnValSerPhe 1600
Db 6593 TGC----- 6595
Qy 1601 ValThrGluGlnAsnTrpAspSerLeuGluValPheAspGlyAlaAspAsnThrValThr 1620
Db 6595 ----- 6595
Qy 1621 MetLeuGlySerPheSerGlyThrThrValProAlaLeuAsnSerThrSerAsnGln 1640
Db 6596 -----ATCCCGTG----- 6604
Qy 1641 LeuTyrLeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHisLeuGluTyr 1660
Db 6604 ----- 6604
Qy 1661 LysThrValGlyLeuSerSerCysProGluProAlaValProSerAsnGlyValLysThr 1680
Db 6605 -----CGATGGGAGAGCTCCACACATCGCAATGGCTTACCCGAGT 6646
Qy 1681 GlyGluArgTyrLeuValAsnAspValValSerPheGlnCysGluProGlyTyrAlaLeu 1700
Db 6647 GGGACAACATACAGTTTGGGCGCTGTGGCTTACACTGCGCAACAGGATTCATATTC 6706
Qy 1701 GlnGlyHisAlaHisIleSerCysMetProGlyThrValArgArgTyrAsnTyrProPro 1720
Db 6707 AAGGGGAGAGAGAGACAGTGT-----GAGGCCACAGACAGTGAAGTAAACCCACG 6760

QY 1721 ProLeuCywilealaglnCywilegylThrValgluGluMetgluGlyValIleLeuSer 1740
 |||
 DB 6761 CCCACCTGC----- 6769
 QY 1741 ProgluPheProgluYanThrProSerAsnMetAspCysSerTrpIleAlaLeuPro 1760
 |||
 DB 6770 -----CATCCT-----GTGCTCTGTAAAGCGACCACTTAAGGTTGAG 6805
 QY 1761 ValgluPhegluAlaHilegluPheLeuAsnPheserThrghuProAsnHileAspTyr 1780
 |||
 DB 6806 AACGGCTTCTGGAGCAACCACTGGACGACCTTTGAGAGCAAGCAAGGTTCCAGTCC 6865
 QY 1781 IlegluIleArgAsnGlyProTyrGluThrSerArgMetMetGlyArgPheserGlySer 1800
 |||
 DB 6866 -----AACCCAGGCTATTAAGCA-----GCCGGAAGT 6892
 QY 1801 GluLeuProSerSerLeuLeuSerThrSerHileGluThrValTyrPheHileSerAsp 1820
 |||
 DB 6893 -----CCTGTGTTGTTGGCAAGCCAAATCGC-----CACTGGCAAGCGAC 6934
 QY 1821 HileSerIleAsnArgProgluPheIleuGluTyrGlnAlaTyrGluLeuGlnGlyCys 1840
 |||
 DB 6935 GCCCTCTGCTGCACCCCTCTC-----AACTGT 6964
 QY 1841 ProAspProgluPhePheAlaAsnGlyIleValArggluAlaGlyTyrAsnValGlyGln 1860
 |||
 DB 6965 GGGAAACCCCTCCATTCAGAAATGCGCTTTTGAAGAGAAACCTTGAAGTGGGCTC 7024
 QY 1861 SerValThrPhegluGlyLeuProgluTyrGlnLeuThrGlyHileProValLeuThrCys 1880
 |||
 DB 7025 AAGGTTCAGTTGCTGTAAATGAGGATATGAGCTCGTGGTGAATATTCCTGGACTTGC 7084
 QY 1881 GlnHilegluThrAsnArgAsnThrAsp---HileProLeuProIleCysGluValPro--- 1898
 |||
 DB 7085 CAG-----AAATCGCAATGGAGTAAAGCAAGCCGAAAGTGT---GTCCCAACC 7135
 QY 1899 ---CysgluGlyAsnIleThrSerSerAsnGlyThrValTyrSerProgluPheProSer 1917
 |||
 DB 7136 AAGGTGCAGAGCTCTCTTGAAGAAACAGCTCGATTTGAAGAA----- 7183
 QY 1918 ProTyrSerSerSerGlnAspCysValTrpLeuIleThrValProIleGlyHileGlyVal 1937
 |||
 DB 7184 -----TTAGCTTCGAGGTGAGGATGATGACATTCCTGTAAAGAGGCGATGCTTG 7237
 QY 1938 ArgLeuAsnLeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTrpAsp 1957
 |||
 DB 7238 CAA---GCCCCCTCTGCTGTAAAGTGTGCTGCAATCCGGGCAA-----TGAAT 7282
 QY 1958 GlyProGlnGlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaIleValThr 1977
 |||
 DB 7283 GGT----- 7285
 QY 1978 ValGlnSerSerSerAsnGlnValIleuLeuIleuPheHileArgAspAlaIleThrGlyGly 1997
 |||
 DB 7285 ----- 7285
 QY 1998 IlePheAlaIleAlaPheSerAlaTyrProLeuThrIle---CysProProPro 2014
 |||
 DB 7286 -----TCCCTTCTTAATTTGTAAAGATGTCTTGTGCTCCCTCCCT 7324
 QY 2015 ThrIleLeuProAsnAlaGluValIleThrGluAsnGluGluPheAsnIleGlyAspIle 2034
 |||
 DB 7325 CCCTTGATCCCTTCCGC---GTCCCTGCGCTTCCGCGTCTTTCATTTTGGAGTACT 7381
 QY 2035 ValArgTyrArgCysLeuProgluPheThrLeuValGlyAsnGluIleLeuThrCysIle 2054
 |||
 DB 7382 GTCAAGATCTGCTGTGCGAGCGGTTTTCTTAAGAGCAAGTCAACATCTCTGCGAG 7441
 QY 2055 LeuGlyThrTyrIleGluIlePheGluGlyProProIleCys---GluValHileCysPro 2073
 |||
 DB 7442 GCTGATAGC-----ACCTGGAAGTTTCTCAATGCGCAATGCGTTCGAGTAAGATGCC 7495
 QY 2074 ThrAsnGluLeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTyrProgluSer 2093

DB 7496 CAACCTAGAGATCTCTCAAC-----GGATCATCTACAGTACAGGCTTGCTATCTC 7549
 QY 2094 TyrProGlnPheGlnThrCysSerTrpLeuValArgValIleuProAspTyrAsnIleSer 2113
 |||
 DB 7550 AGACCAACGCTTACACTGC-----AAGCCAGGCTTTGAGTTA--- 7558
 QY 2114 LeuThrValGluTyrPheLeuSerGluIleGlyThrAspIleuPheGluIlePheAspGly 2133
 |||
 DB 7589 -----GTGGGCAATGCTACACCTCTGTGGG 7615
 QY 2134 ProSerGlyGln-----SerProLeuLeuIleAlaLeuSerGlyAsnTyrSer 2149
 |||
 DB 7616 GAAATGCGCAGTGGCTCGAGAGAAACCAATGCGAAACCATTTGAA----- 7663
 QY 2150 AlaProLeuIleValThrSerSerSerAsnSerValTyrIleuArgTrpSerSerAspHis 2169
 |||
 DB 7663 ----- 7663
 QY 2170 AlaTyrAsnArgIleGlyPheIleArgTyrSerAlaProTyrCysSerLeuProArg 2189
 |||
 DB 7664 -----TCCCAAGGCCAAG 7678
 QY 2190 AlaProLeuHilegluPheIleLeuGlyGlnThrSerThrGlnProgluGlySerIleHis 2209
 |||
 DB 7679 GAGATTTTAATAGC---CAATCTCTTCGCTGAGCTTTCAGTATGAGCAAAACATCACA 7735
 QY 2210 PheGlyCysAsnAlaGlyTyrArgLeuValGlyHileSerMetAlaIleCysThrArgHis 2229
 |||
 DB 7736 TACTTTGTATCCGGGGCTTCCGGCTCGAAGGTGCCAAATCCCTGACCTGTTAAGACAA 7795
 QY 2230 ProGlnGlyThrIleuLeuTrpSerGluAlaIleProLeuCysGlnAlaLeuSerCysGly 2249
 |||
 DB 7796 GGTGAC-----TGGATATGAGATCCCTCTTGTATGATGCATCTGACAGT 7843
 QY 2250 LeuProGluAlaProIleAsnGlyMetValPheGlyIleGlyTyrThrValGlyThrIle 2269
 |||
 DB 7844 GACCCACAGCCCAATGAATAATGTTGTGTGAAGGTGCGGATTAACAATGCGGCGCATG 7903
 QY 2270 AlaValTyrSerCysSerGluGlyTyrHileLeuGlnAlaGlyAlaIleThrAlaGlu 2289
 |||
 DB 7904 ATCATTAATAGCTGCTTCCCTGGGTTTCAGAGTGTGTCATGCAATGACAGACAC- 7957
 QY 2290 CysLeuAspThrGlyLeuTrpSerAsnArgAsnValProProGlnCysValProValThr 2309
 |||
 DB 7958 TGTGAAGATCGGGA---TGTCAAGCTCCAGC---CCAACTGTGTACCATTAAGAC 8008
 QY 2310 Cys-----ProAspVal-----SerSerIleSerValGluHileGlyArgTrp 2323
 |||
 DB 8009 TGGGCTCCCTCTCTCATATAGACTTTGGTGAATCTGTAAAGTCAAGATGGCCAGGAA 8068
 QY 2324 ArgLeuIlePheGlu----- 2328
 DB 8069 CATTTGATCAAGAAATGACATGATGAGATGCCAATATCTGCTCACCTCAACATTTG 8128
 |||
 QY 2329 -----ThrGln 2330
 DB 8129 GAAGCAACAGCTAAGCCTTGAAATATCAAAAGAGTGGCTGCTTCATGATGATCCAC 8188
 QY 2331 TyrGlnPheGlnAlaGlnLeuMetLeuIleCysAspProgluTyrTyrTyrThrGlyGln 2350
 |||
 DB 8189 TTCTCTATATGACAGATGTTTCTCTACAGCTGCGAGCTGTTTGAATGCTGGGAAATC 8248
 QY 2351 ArgValIleArgCysGlnAlaAsnGlyIleTrpSerLeuGlyAspSerThrProThrCys 2370
 |||
 DB 8249 CCGTGTGATCTGCGCAGAAATGTGATCGTAAT-----GTAACCGACCCCTTTCG 8302
 QY 2371 ArgIleIleSerCysgluGluLeuProIle---ProProAsnGlyHileAspGlyThr 2389
 |||
 DB 8303 ATTTCATTAAGT---GATTTGCTGTGTCTCCGAAATAGGCTTTTACATTTGACA 8359
 QY 2390 LeuSerValTyrGlyAlaThrAlaIlePheSerCysAsnSerGlyTyrThrLeuValGly 2409
 |||

Db 8360 CAGACGACTATGGGACAGTGTGACAAATATAGCTGCAAGCCGGGGCACTTTTGAAGAGC 8419
Qy 2410 SerArgValArgGluCysMetAlaSerGlyLeuTrpSerGlySerGluValArgCysLeu 2429
Db 8420 TCCCACTTAAGACTGTCTGTGCAAGAAATAGACAGTGGAGTGCACCTGTTCCACCTGTGA 8479
Qy 2430 AlGlyHisCysGlyThrProGluProIleValAsnGlyHisIleAsnGlyGluAsnTrp 2449
Db 8480 GCCATCTCATGCAATGACCAACCACTCTGGAAATGATCCATCAAGAGATGATGATAC 8539
Qy 2450 SerTyrArgGlySerValValTyrGlnCysAsnAlaGlyPheArgLeuIleGlyMetSer 2469
Db 8540 TCTTACCTGGGTGTGTACTAGAGTGTGACTCTGGTATATTCATATGCTCTGTAAG 8599
Qy 2470 ValArgIleCysGlnGlnAspHisIleTrpSerGlyLysThrProPheCysValProIle 2489
Db 8600 AAGAGCATGCAAGAAATAGAGATTGGATGGGATGAGCCCATGTGTATTCCTGTA 8659
Qy 2490 ThrCysGlyHisProGlyAsnProValAsnGlyLeuThrGlnGlyAsnGlnPheAsnLeu 2509
Db 8660 GACTGTGGCTCACCCCAAGTCCCAACCAATGGCCGAGTGAAGAGAAATACACATTC 8719
Qy 2510 AsnAspValValArgPheValCysAsnProGlyTyrMetAlaGluGlyAlaArgSer 2529
Db 8720 CAAAGGAGATACATCTCTGTGCGGTGAAGGGTTACTGGAAGAGCCAGAGTGT 8779
Qy 2530 GlnCysLeuAlaSerGlyGlnTrpSerAspMetLeuProThrCysArgIleIleAsnCys 2549
Db 8780 ATGTGCTTACCAATGAGATTGAGTGTGCACCTCCAGCTGCATGCTGTTAAGATG 8839
Qy 2550 ThrAspProGlyHisGlnGlnAsnSerValArgGlnValHisAlaSerGlyProHisArg 2569
Db 8840 CCGCCCAACACAGAGTCCCAATGGGTG-----GCAATGGCCCTAAGAC--- 8884
Qy 2570 PheSerPheGlyThrThrValSerTyrArgCysAsnHisGlyPheTyrLeuGlyThr 2589
Db 8885 TATGGTTTCAAGAAAGAAAGTACGCTTCACTGTCTAGAGGCTATGTGCGAGGGGGCT 8944
Qy 2590 ProValLeuSerCysGlnGlnAspGlyThrTrpAspArgProArgProGlnCysLeuLeu 2609
Db 8945 CCAAGACTCACCTGATGCCAATGAGGACTGGAGTGGAGACAGAGTCCCTGTCTGAACCA 9004
Qy 2610 ValSerCysGlyHisProGlySerProProHisSerGlnMetSerGlyAspSerTyrThr 2629
Db 9005 GCTACCTGTGTCTCTCTGCGCAACCTTCCCAAGGCTCCCTAATGGCTTTCTTTTAT 9064
Qy 2630 ValGlyAlaValArgTyrSerCysIleGlyLysArgThrLeuValGlyAsnSerThr 2649
Db 9065 CATGGGGGCAACATACAGTATCAGTGTCTTACTGTTATTAAGCTTCATGGAACCATCA 9124
Qy 2650 ArgMetCysGlyLeuAspGlyHisIleTrpThrGlySerLeuProHisCys----- 2665
Db 9125 AGAAGATGCTTCTCCCAATGATGCTCTGAGCGGCACTCGCCATCTGCTTACCTTGCAAG 9184
Qy 2666 -----SerGlyThrSerAlaGlyVal----- 2672
Db 9185 TGTTCACACCCATCATTTCAACAGGAACATCAAGCAACTATTGGATGTGGAAG 9244
Qy 2672 ----- 2672
Db 9245 ACGGTCAGATTGATGCTTCAAGGCTTCAGCTGCTTGAATTCGAAATCACCTGT 9304
Qy 2673 -----CysGlyAsp----- 2675
Db 9305 GATGCCAATGGCCAAATGTCTGACGTCCACTGTGTGAGCAGCTGAGTGGGGCTCTC 9364
Qy 2676 ProGlyIleProAlaHisGlyIleArgLeuGlyAspSerPheAspProGlyThrValMet 2695
Db 9365 CCAACCATACC---AAGCAATTGTCTCTGAGGCAAGCTTTCGAGAGCAATGTGTGA 9421
Qy 2696 ArgPheSerCysGluAlaGlyHisValLeuArgGlySerSerGluArgThrCysGlnAla 2715
Db 9422 ACTTACAGTGTGACAGCTGTGCTACACCATGCAAGTAGTTCAGATTGTATGAGAA 9481

Qy 2716 AsnGlySerTrpSerGlySerGlnProGluCysGlyValIleSerCysGlyAsnProGly 2735
Db 9482 AAGCAATATGAGCGACAGCTTACCCACAGTGTGAACCCCTGTCTGTGACCCCAACA 9541
Qy 2736 ThrProSerAsnAla----- 2740
Db 9542 ACTGTAGCCAAATGCACTGGCAACAGAGAGGCTCATCTATGAAGCAAAATGAAATC 9601
Qy 2740 ----- 2740
Db 9602 AGGTGTGGAAGGATATGATGATTCGATACATACATTCACCTGCAGCAAGAT 9661
Qy 2740 ----- 2740
Db 9662 GGCATTTGGTCCCTGAAAGAAATCACCTGCAGTCTTAAAAATGCCCTGTCCATCCAAC 9721
Qy 2741 -----ArgValAlaPhe-----SerAspGlyLeuValPheSerSerIleValTyrGlu 2757
Db 9722 ATGACACGATACGTTTTCACGAGATGACTTCCAGTGAACAGACAAAGTTTCTGTCTCA 9781
Qy 2758 CysArgGluGlyTyrTyrAlaThrGlyLeuLeuSerArgHisCysSerValAsnGlyThr 2777
Db 9782 TGTGCAGAAAGGTTTAAACCAAGAGAGTGAATGCTGTCACATGCCCCAGCGGTACA 9841
Qy 2778 TrpThrGly-----SerAspProGluCysLeuValIleAsnCysGlyAspProGlyIle 2795
Db 9842 TGGAGCCACCAATTTCTGATGAATCTGTATCCACAGTTGTTGTGGGATCCTGAAGC 9901
Qy 2796 ProAlaAsnGlyLeuArgLeuGlyAsnAspPheArgTyrAsnIleThrValThrTyrGln 2815
Db 9902 CCAAGCGCATAGCTCGGTGTGGCAATAAGACAGCTTTGGAAGCAACATTTGTTACAG 9961
Qy 2816 CysValProGlyTyrMetCysLeuSerHisArgValSerValLeuSerCysThrLysAsp 2835
Db 9962 TGTGACCTCGCTTACAAATTAAGAGGGGAACAGAGAACGAAATC-----TGCCAGAGAAAC 10015
Qy 2836 ArgThrTrpAsnGlyThrLysProValCysValAlaLeuMetCysLysProProLeu 2855
Db 10016 AAGACGTGAGTGAAGAGGTGGCACTGTGCAAGAAACAGATGTGAGCTCCAGCTGAG 10075
Qy 2856 IleProAsnGlyLysValValGlySerAspPheMetTrpGlySerSerValThrTyrAla 2875
Db 10076 TTTCCCAATGGGAAGGCTGTCTTGAAGAAACACACATCTGCAGCCAGCTTCTGTTTCC 10135
Qy 2876 CysLeuGluGlyTyrGlnLeuSerLeuProAlaValPheThrCysGluGlyAsnGlySer 2895
Db 10136 TGTCAAGAGGCTTACACCTGGAAGGATCCCCGAGGACACACTGCACCTCAATGGAAC 10195
Qy 2896 TrpThrGlyLeuLeuProGlnCysPheProValPheCysGlyAspProGlyVal---Pro 2914
Db 10196 TGAATACCTGACTTCCCTCTGCAACCAATTCATGATCCCTGCTTTGTGATTCCT 10255
Qy 2915 SerArgGlyArgArgGluAspArgGlyPheSerTyrArgSerSerValSerPheSerCys 2934
Db 10256 GAGAAAGCGGCTTCTTGAAGAAAGATTATGTGCACCAAGATGATCATCAAGTGC 10315
Qy 2935 HisProProLeuValLeuValGlySerProArgArgPheCysGlnSerAspGlyThrTrp 2954
Db 10316 AAGGAAGGCTTCTCTCAAGAGCAATGTGTATCATCACTGCAGCCCTGACAGACATGG 10375
Qy 2955 SerGlyThrGlnProSerCysIleAspProThrLeuThrCysAlaAspProGlyVal 2974
Db 10376 ACGCACCAATGCCAGATG-----GAAAAATCTCTGTGTGCTCTCAAGTCAAC 10426
Qy 2975 ProGlnPheGlyIleGlnAsnAsnSerGlnGlyTyrGlnValGlySerThrValLeuPhe 2994
Db 10427 GTAAGAAATGCAATTGCTCGAGAGTGTAT---TACAGTATAGGGAGACATGATCACTAC 10483
Qy 2995 ArgCysGlnLysGlyTyrLeuLeuGlnGlySerThrThrArgThrCysLeuProAsnLeu 3014
Db 10484 TCTGTGTACAGTGTGCTACATGCTGAAGAGTTCCTCCGAGAGTGTGCTTGAAGAAATGGA 10543

QY 520 ----- 520
 Db 2411 GAGATGTTCTAACAAGCAGCTCGTTGTGATGACACAGATCTGATGAGAAGTTTCTGAA 2470
 QY 521 -----ThrlleuserProglYlYrProglYlYrAspAla----- 534
 Db 2471 GCATTGAGACGACCTGGAAAAATGGTCCCATCTTTGTAGTATGACAGAGGACATT 2530
 QY 535 -----LeuserCyAlaIrrValIleGluAlaGlnProglYlYr 547
 Db 2531 GACTGAGACTGGAGAGAACCTGACCAAAAATATTGGCTGAGAAATGATTAATGACTAT 2590
 QY 548 Pro-----IleuSlethr 552
 Db 2591 GAAATGGCTTTCATTTGACCAAGTGGCTGGGGTGCAGCTAATAGCTGACTTCTCT 2650
 QY 553 PheAspArgPheIleThrGluValIleuValAspThrLeuGluValAlaArgAspGlyArg-- 571
 Db 2651 TAGATGACTTCTCGACACTGTGCAAGAAACAGCAAGCATGGGCAATGCCAAGTCC 2710
 QY 571 ----- 571
 Db 2711 TCACGATTAAAGAGTCCCATTTATCTGACTATTAATTAAGTTAATTTTAACTC 2770
 QY 572 ThrTySerAlaProIleuIleGlyValIleGlyThr----- 584
 Db 2771 ACAGTATGTGTGCTATTCATCCGATGAAGAAATGATACCTTGAAATGGAAAAATCAGCA 2830
 QY 585 GluValProGlnPheLeuIleSerThrSerAsnTyIleuTyIleuPheSerThrAsp 604
 Db 2831 CGACTCTTCAGACATGGAACATGACAAATTAATGAAAGAACTGCACAAAGAC 2890
 QY 605 LysSerHisSerAspIleGlyPheGlnLeuAlaGlyThrIle----- 619
 Db 2891 CCCATGATTTCC-----TTTCAGCTTGACATCAAAATACTTATAGCCGACAGCAAT 2941
 QY 620 -----ThrlleuGlnSerAspHisCysLeuAspProglYlYrProValIleuGlyGln 636
 Db 2942 TCATTGAAACAAAAAAGGCTTCCCTTGCAGACAGGCTCAGTGCAGAGAGGCGCT 3001
 QY 637 ArgHisIleGlyAsn-----AspPheTyValIleValIleuValIlePheSerCys 652
 Db 3002 ATGTGTCAATTGCCCTTGGGAACTATTAAT-----AATCGGAACATTTCCACTGT 3055
 QY 653 AspSer-----GlyTyThrLeuSerAspIleGlyProIleuGluCysGlyAla-- 667
 Db 3056 GAAAGCTGCCGAGATCGATCTTATCAAGATGAAAGAGGCA--CTTGAGTGCAGCTT 3112
 QY 668 -----ProAsnPheGlnIrrPheArgAlaLeuProSerCysGlyAla 681
 Db 3113 TGCCCTCTGGGATGTACAGGAATATATCATTCAGAAACATCTGTGATGTAAAGCT 3172
 QY 682 LeuCyGly----- 684
 Db 3173 CAGGTGAACAAAGCAGCTACTATACAGTGAAGCTTGAGTGTGATCTGCTCACTG 3232
 QY 685 GlyIleIleGln-----GlySerSerCylYrIleLeuSerPro----- 697
 Db 3233 GGCATTTATCAGCAAAATTTGGTTCCGAGAGCTGCTCTGTCTCAAGAAAACCTCA 3292
 QY 698 -----GlyPhePro----- 700
 Db 3293 ACTGTGAAGAGAGAGCGGTGACATTTCTGATGTGAGATCTCTTGCCAGAGAGAAA 3352
 QY 701 -----AspPheTyProAsnAsn 706
 Db 3353 TTCTGCGCTTGGGTTAATGCTCTGTCAACCATGTCTGTGACTATTAACCACTAAT 3412
 QY 707 -----LeuAsnCyThrIrrIleIleGluThrSerHisIleGlyValGly 720
 Db 3413 GCAGGAAAGGCTTCTGTGCTGCTGTCTCTT-----TATGAAACATACC 3457
 QY 721 ValPhePheThrPheHisIleuGlnSerGlyHisAspTyIleuLeuIleThr 740

Db 3458 CCATTGCGTGGTCCAGATCC-----ATCACA 3484
 QY 741 GluAsnGlySerPheThrGlnProIleuArgGlnLeuThrGlySerArgLeuProAlaPro 760
 Db 3485 GAATGTTCAAGTTTATGTTCAACTTCTCACGGCAGAGAAAGTGTGTGCC--CCT 3541
 QY 761 IleSerAlaGly----- 764
 Db 3542 GCCTTCTTGGACATATTAAAAAGAGCATGAATACAGACTCAGCTTTCCATGATGC 3601
 QY 765 -----LeuTyGlyAsnPheThrAlaGlnVal-----ArgPheIleSer 777
 Db 3602 TTCTTTAACCTTGGCCACAAATAGTGAACCTGCAGCAACTTGGCGCTGTATGTGTGT 3661
 QY 778 AspPheSerMetSerTyIleGlyIlePheAsnIleThrPheSerGluTyAspLeuGluPro 797
 Db 3662 CTCTGTCCACTTGGATATACAGGCTTAAAG-----TGTGAAACAGACATCGATGAG 3712
 QY 798 CysGluGluProGluValProAlaIleTySerIleArgIleGlyLeuGlnPheGlyValGly 817
 Db 3713 TGCAGC-----CCACTGCTTGCTCAACATGAGAGTTGTAAAGACTAGTT 3760
 QY 818 AspThrIleuThrPheSerCysPheProglYlYrArg----- 829
 Db 3761 GGGAAATTCATTTGTGAGTCCCATCAGTTTACACAGGTGACGGGTGTGAAGAAATATA 3820
 QY 830 -----LeuGluGlyThrAla----- 834
 Db 3821 AATGAGTGTAGCTCAGTCTTGTTTAAATGAAGAAATCTGTGTTAGTGTGTGCTGGC 3880
 QY 835 ---ArgIleThrCysLeuGlyGly-----ArgArgArg 844
 Db 3881 TATGTTGACATGtGTGAAGAGATTGTAGGCTGCATGTGAAACAGAACTCAATGA 3940
 QY 845 LeuTrpSerPro----- 849
 Db 3941 TCCAGTCAAAACCATGCTTAATATGACATGTGTGAAGACCAAGTTGGGGAAATTTCTG 4000
 QY 850 -----LeuProArgCysVal 854
 Db 4001 TGCAAATGCCCACTCGATTTTGGGTACCCGATGTGGAAGAAAGATGATGATGCTCAT 4060
 QY 855 AlaGlu---CysGlyAsnSerValIleGlyThrIleGlyThr-----Leu 868
 Db 4061 AGTCAGCATGCAAAATATGAGACTACCTGTAAAGACGGTGCATATGCTTCAAGTGCCTG 4120
 QY 869 LeuSerProAsnPhe-----ProValAsnTyAsn----- 878
 Db 4121 TGTGAGCTGGCTTACAGAGATCAACATGTGAATGAAATGAAATGATGCTCATAT 4180
 QY 879 -----AsnAsnHisIleGlyCysIle-----TySerIleGlnIleThrGln 890
 Db 4181 CCATGTGAATATCAGGCCACCTGTGTGATGAATTAATTCATACAGTTGTAATGTCTAG 4240
 QY 891 ProGly-----LysGlyIleGlnLeuValAla 899
 Db 4241 CAGGATTTTACAGCAAAAGGTGTGAACAGAACAGCTTACAGGCTTTAACTGAT-- 4297
 QY 900 ArgAlaPheGluLeuSer----- 905
 Db 4298 -----TTTGAAGTTTCTGGCATTTATGATATGTCAATGCTAGATGCTCCATCT 4351
 QY 905 ----- 905
 Db 4352 CTCGATGCTTAACCTGTACTCTTGTGAGTGAATCTGTGACGACATGAATATGGAACA 4411
 QY 906 -----Glu 906
 Db 4412 CCAATCTCTTATGAGCTGTGATGAACGGACAGCAATACCTTGTCTCTGACTGATTAAC 4471
 QY 907 GlyAspValLeu----- 910


```

QY 1496 SerValProGluProArgTyrGly-----LysArgLeuGlySerAspPheSerVal 1512
Db 6263 AAACCTCCATCGCTTCCCTATAGCATCTTGATCTGTGAGCAAAATTTGCACCT 6332
QY 1513 GlyAlaIleValArgPheGluCysAsnSerGlyTyrAlaLeuGlnGlySerProGluIle 1532
Db 6323 GCGTCAGTTGAGCTTTAAATGATGAAAGCTTTGTAAGTGAACCTCAGCAAAAGATT 6382
QY 1533 GluCysLeuProValProGlyAlaLeuAlaGlnTyrAsnValSerAlaProThrCysVal 1552
Db 6383 GAATGATG-----AGAGGTGGGCGAGTGAAC----- 6409
QY 1553 ValProCysGlyGlyAsnLeuThrGluArgGlyThrIleLeuSerProGlyPhePro 1572
Db 6410 -----CCT 6412
QY 1573 GluProTyrLeuAsnSerLeuAsnCysValTyrPylsIleValValProGluGlyAlaGly 1592
Db 6413 TCCCCCATG-----TCCATCCAGTGC----- 6433
QY 1593 IleGlnIleGlnValIleValSerPheValThrGluGlnAsnTyrAspSerLeuGluValPhe 1612
Db 6433 ----- 6433
QY 1613 AspGlyAlaAspAsnThrValThrMetLeuGlySerPheSerGlyThrThrValProAla 1632
Db 6434 -----ATCCCTGTG 6442
QY 1633 LeuLeuAsnSerThrSerAsnGlnLeuTyrLeuHisPheTyrSerAspIleSerValSer 1652
Db 6442 ----- 6442
QY 1653 AlaAlaGlyPheHisLeuGluTyrLeuThrValGlyLeuSerSerCysProGluProAla 1672
Db 6443 -----CCGTGTGAGAGCCACCA 6460
QY 1673 ValProSerAsnGlyValIleValThrGlyGluArgTyrLeuValAsnAspValIleSerPhe 1692
Db 6461 AGCATCATGAAATGCTATGCAAGAGATCAAACTACAGTTTGGAGCCATGCTGGCTTAC 6520
QY 1693 GlnCysGluProGlyTyrAlaLeuGlnGlyHisAlaHisIleSerCysMetProGlyThr 1712
Db 6521 AGCTGCAACAAAGGGTCTTCATCAAAAGGAAAGAGACACCTGC-----GAAGCC 6574
QY 1713 ValArgTyrPheAsnTyrProProProLeuGlyIleAlaGlnCysGlyGlyThrValGlu 1732
Db 6575 ACAGGCAAGTGAAGTACCTATACCGACGTGC----- 6607
QY 1733 GluMetGluGlyValIleLeuSerProGlyPheProGlyAsnTyrProSerAsnMetAsp 1752
Db 6607 ----- 6607
QY 1753 CysSerTyrPylsIleAlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPhe 1772
Db 6608 -----CACCGGTATCTTGGT----- 6625
QY 1773 SerThrGluProAsnHisAspTyrIleGlnIleArgAsnGlyProTyrGluThrSerArg 1792
Db 6626 -----GAACCACCT-----AAGTTGAAGATGGCTTTCTGAG----- 6658
QY 1793 MetMetGlyArgPheSerGlySerGluLeuProSerSerLeuLeuSerThrSerHisGlu 1812
Db 6659 -----CATACA 6664
QY 1813 ThrThrValTyrPheHisSerAsp---HisSerGlnAsnArgProGlyPheLys---Leu 1830
Db 6665 ACTGCGCAGATCTTTGAGGTGAAGTGAAGTATCATGTAACCGGCGATATAAGTCAATGC 6724
QY 1831 GluTyrGlnAlaTyrGluLeuGln----- 1838
Db 6725 GGAAGTCTGTATTTGTCTGCCAAGCAATCGCCACTGSCACAGTAATCCCTCTGATG 6784
QY 1839 -----GluCysProAspProGluProPheAlaAsnGlyIleValArgGlyAla 1854

```

```

Db 6785 TGTGTTCTCTCCACTGTGGAAAACTCCCGATCCAGAAATGGCTTCAAGAGAGA 6844
QY 1855 GlyTyrAsnValGlyGlnSerValThrPheGluCysLeuProGlyTyrGlnLeuThrGly 1874
Db 6845 AACTTGAAATAGAGGTCCAAGGTCATCTTTCTGTATAGAGGTTATAGCTTGTGT 6904
QY 1875 HisProValLeuThrCysGlnHisGlyThrAsnArgAsnTyrAspHisProLeu---Pro 1893
Db 6905 GACAGTTCTTGACATGTCAG-----AAATCGGAATGGAATAGAAGTCAATCCA 6958
QY 1894 LysCys-----GluValProCysGlyGlyAsn----- 1902
Db 6959 AAGTGCATCCCTGCCAAGTCCAGACCGCCCTCTTGGAACACAGCTAGTATTAAG 7018
QY 1903 ---IleThrSerSerAsnGlyThrValTyrSerProGlyPheProSerProTyrSerSer 1921
Db 7019 GAGTTGACCAACGAGGTAGAGTGTG-----ACATTTCTCTGT 7057
QY 1922 SerGlnAspCysValTyrPheLeuIleThrValProIleGlyHisGlyValArgLeuAsnLeu 1941
Db 7058 AAGAA-----GGGCAT---GTCCTGCAAGGCCCC 7084
QY 1942 SerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTyrAspGlyProGlnGln 1961
Db 7085 TCTGTCTGAATAGCTTGCCATCCACGA-----TGGAAAT----- 7120
QY 1962 ThrAlaProArgLeuGlyValPheThrArgSerMetAlaLysLysThrValGlnSerSer 1981
Db 7120 ----- 7120
QY 1982 SerAsnGlnValLeuLeuLysPheHisArgAspAlaIleThrGlyIlePheAlaIle 2001
Db 7120 ----- 7120
QY 2002 AlaPheSerAlaTyrProLeuThrLys-----CysProProThrIleLeuPro 2018
Db 7121 -----GACTCTTCCCTGTTGTAAGATTGTTCTTGTACCACCACTCCCTTAAT--- 7171
QY 2019 AsnAlaGluValValThrGlnAsnGluGlnPheAsnIleGlyAspIleValArgTyrArg 2038
Db 7172 TCCTTTGGTGTCCCATCTCTTCTGCTCTTCAATTTTGAAGTACGTCAAGTATCT 7231
QY 2039 CysLeuProGlyPheThrLeuValGlyAsnGlnIleLeuThrCysLysLeu---GlyThr 2057
Db 7232 TGTGAGTGGGTTTCTTAAGAGAAATTTCAACACCTCTGCCAAGCTGAGGCACC 7291
QY 2058 TyrLeuGlnPheGluGlyProProIleCys---GluValHisCysProThrAsnGlu 2076
Db 7292 TGG-----ACCTCCACCTGCCAAGATGTGTCAGTGAAGATCTCCCAACCTGAG 7342
QY 2077 LeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTyrProGlySerTyrProGln 2096
Db 7343 GAATTCGCCAAT-----GGAATCATGATGTGCAAGCCCTTGCTTCCAGCAGCAGCT 7396
QY 2097 PheGlnThrCysSerTyrPheValArgValGlnProAspTyrAsnIleSerLeuThrVal 2116
Db 7397 CTCTATACCTGC-----AAGCAGGCTTGAATG----- 7426
QY 2117 GluTyrPheLeuSerGluLysGlnTyrAspGluPheGlnIlePheAspGlyProSerGly 2136
Db 7427 -----GTGCGA 7432
QY 2137 GlnSerProLeuLeuLysAlaLeuSerGlyAsnTyrSerAlaProLeuIleValThrSer 2156
Db 7433 AATCTACCAACCTTTTGAGAAATGGTCACTGC----- 7468
QY 2157 SerSerAsnSerValTyrLeuArgTyrSerSerAspHisAlaTyrAsnArgLysGlyPhe 2176
Db 7469 -----CTTGAGCA 7477
QY 2177 LysIleArgTyrSerAlaProTyrCysSerLeuProArgAlaProLeuHisGlyPheIle 2196

```



```

Db      7478 AACCAACATGTAAGCCATTGAGTGCCTGAAACCAAGAGATTTGTAATGCG---AAA 7534
Qy      2197 LeuclglnThrSerThrInglProglYglYserIleHisPheglYcysAsnAlaIglYThr 2216
Db      7535 TTCTCTTAACAGCAAGCTTAACACTATGAGACAGACCGTCTACTGTCGACCGAGGCTTT 7594
Qy      2217 ArgLeuValIglYHisSerMetAlaIleCysThrArgHisProglnglYThrIleLeuThr 2236
Db      7595 CGGCTCGAAGGCTCCACAGTGCCTTAACCTGTTTAGACAGAGTGAAT-----TGG 7642
Qy      2237 SerclAlaIlePheLeuCysGlnAlaLeuSerCysglYLeuProgluAlaProLYasn 2256
Db      7643 GATGTAGATGCCCATCTTGCAATGCCATCCACTGATGATCCCAACAACCATTTGAAAT 7702
Qy      2257 GlYMetValPheglYglYglYThrValIglYThrIysAlaValIYrSerCysSercln 2276
Db      7703 GGTTTTGAGAAAGTGACGATTACAGCTATGTCCTCATATCATCTACATGATGCTTCCCT 7762
Qy      2277 GlYThrIleLeuGlnAlaIglYAlaIglYAlaThrIleIleCysLeuAspThrIglYLeuThr 2296
Db      7763 GGGTT-----CAGGTGGCTGTCGTCATGCCATGCAAGCTGTGAAGTCAAGA---TGG 7813
Qy      2297 SerAsnArgAsnValProProglncYsValProValThrCys----- 2310
Db      7814 TCAGATTCC-----ATCCCAACATGTATGCCAATAGACTGTGGCTCCCTCATATATA 7867
Qy      2311 -----ProAspValSerSerIleSerValIglYHisglY----- 2321
Db      7868 GATTGTGAGACTGTACTTAACTCAAGAGATGACAGAGATTTTGTGACAAAGAACGAC 7927
Qy      2321 ----- 2321
Db      7928 ATGATGGAAGTTCATATGTGATCTCACCCTCTTATCATTTGGAGACAGTGGCTAAA 7987
Qy      2322 ArgTrpArgLeuIlePheglY-----ThrGlnYrGlnPheGlnAla 2335
Db      7988 ACCTGGGAAATAATCAAGAGAGTCTCTGCTACATTCATCAAACTTTCTGTATGTAAC 8047
Qy      2336 GluMetLeuIleCysAspProglYThrYrYrThrIglYAlaIleArgCys 2355
Db      8048 ATGGTTTCATACACCTGTAACTCAAGATGATGAACCTTGGGAAACCTGTGCTATGTC 8107
Qy      2356 GlnAlaAsnGlyIleYrSerLeuGlyAspSerThrProThrCysArgIleIleSerCys 2375
Db      8108 CAGGAAGATGAGACTTGGAAT-----GGCAGTGCACCATCTCTCATTTCAATTGAATGT 8161
Qy      2376 GlyIleuLeuProIleProAsnGlyHisArgIleGlyThrLeuSerValYrGlyAla 2395
Db      8162 GACTTGCCTACTGCTCTGAAATGGCTTTTGGGTTTACAGAGACTAGCATGGGAAGT 8221
Qy      2396 ThrAlaIlePheSerCysAsnSerGlyYrThrIleuValIglYSerArgValArgGluCys 2415
Db      8222 GCTGTGCAATGACTGTAACTGAACTGACACATCTTAGTGGGCTGTGAATGAGCTTTGT 8281
Qy      2416 MetAlaAsnGlyLeuYrSerGlySerclYValIArgCysLeuAlaIglYHisCysglYThr 2435
Db      8282 CTAAAGAAATGAAGAGTGGAGTGGTGCCTCCCAAGCTGTGAAGCATTTCTGCAAAAAG 8341
Qy      2436 ProgluProIleValIleAsnGlyHisIleAsnGlyIleuAsnYrSerYrArgGlySerVal 2455
Db      8342 CCAATATCCAGATGATGATGATCATCAAGAGAAAGCAATCAATACCTGAGACGTTG 8401
Qy      2456 ValYrGlnCysAsnAlaIglYPheArgLeuIleGlyMetSerValArgIleCysGlnGln 2475
Db      8402 TACTATAGTGTGACCCCGATATGTGTGAATGGCACTGAGAGAGAAACATGGCAGAT 8461
Qy      2476 AspHisIleYrSerGlyYrLeuThrProPheCysValProIleThrCysGlnYHisProglY 2495
Db      8462 GACAAAACCTGGATGAGATGAGACCCATTGTCATCTGTGACATGCACTGCCACCCCA 8521
Qy      2496 AsnProValIleAsnGlyLeuThrGlnGlyAsnGlnPheAsnLeuAsnAspValValIysPhe 2515
Db      8522 GTCTCAGCCAAATGCGCAGGTGAGAGAGACGATACATTCCAAAAAAGAGATTGAATAC 8581

```

```

Qy      2516 ValCysAsnProglYrYrMetAlaIglYAlaIleArgSerGlnCysLeuAlaSerGly 2535
Db      8582 ACTTGCAATGAAGGTTCTTCTGTTGAGAGAGCCAGAGTGGGTTTCTTCCCAATGA 8641
Qy      2536 GlnTrpSerAspMetLeuProThrCysArgIleIleAsnCysThrAspProglYHisGln 2555
Db      8642 AGTTGAGATGAGCACTCCCACTGTGCTCGTCAATGTGTCACCCGCCCAACTG 8701
Qy      2556 GluAsnSerValArgGlnValHisAlaSerGlyProHisArgPheSerPheglYThrThr 2575
Db      8702 GCCAATGGGGTGACGAA-----GGCCTGACATATGGCTTCATGAAGAA 8746
Qy      2576 ValSerYrArgCysAsnHisglYPheYrLeuLeuGlyIleThrProValLeuSerCysGln 2595
Db      8747 GTACATTCACACTGTCTATGAGGGCTATACATTGTCAGCGTGTCCAAAATCCTGTCAG 8806
Qy      2596 GlYAspGlyThrTrpAspArgProArgProglncYsLeuLeuValSerCysglYHisPro 2615
Db      8807 TCAGATGCAACTGGATGAGAGATTCCTCTGTAAACCAAGTCAACTGTGACCTCTCT 8866
Qy      2616 GlySerProProHisSerGlnMetSerGlyAspSerYrThrValIglYAlaValArg 2635
Db      8867 GAAGATCTTGCCCATGTGTTCCCTTAATGATTTTCTTATTCATGGGGGCCATATACAG 8926
Qy      2636 TyrSerCysIleGlyIysArgThrLeuValIglYAsnSerThrArgMetCys----- 2652
Db      8927 TATCAGTCTTTCTGCTTATTAAGCTTCATGAATTCATCAAGAGTGTCTTCCAT 8986
Qy      2652 ----- 2652
Db      8987 GGCCTCGAGTGGACAGCTACCTTCTGCTGCTGCTGAGATGTTCCACACAGTAAT 9046
Qy      2652 ----- 2652
Db      9047 GAATATGGAACGTGCAATGGAACAGATTTTGACTGTGAAAGCAGCCCGAATTCAGTGC 9106
Qy      2653 -----GlyLeu-----AspGlyHisIleThr 2658
Db      9107 TTCAAAGGCTTCAAGCTCTTAAGACTTTTGTGAATATCACTGTGAAGCCGATGGCAATGG 9166
Qy      2659 ThrGlySerLeuProHisCysSerGlyYrThrSerValGlyValCysglYAsp---ProglY 2677
Db      9167 AGCTCGGGTTCCTCCCACTGTGAACACATCTCT-----TGTGTTCTCTTCCATG 9217
Qy      2678 IleProAlaHisglYIleArgLeuGlyAspSerPheAspProglYrYrValMetArgPhe 2697
Db      9218 ATACCAATGCGTTTCATCAGTGAAGACCAAGCTCTTGAAGGAA---AATGTGATTAATTAC 9274
Qy      2698 SerCysGlnAlaIglYHisValIleuArgGlySerSerGlnYrThrCysGlnAlaAsnGly 2717
Db      9275 AGCTCAGGTGTGATATGTATCAATCAAGGCAAGCTTCAAGATGATTTGTATACAGAAAGGG 9334
Qy      2718 SerTrpSerGlySerGlnProglYcYsGlyValIleSerCysglYAsnProglYThrPro 2737
Db      9335 GTATGAGCAAGCTTATCCAGTCTGTGAACCCCTGTGCTGTGGGTCCCAAGCGTCTGC 9394
Qy      2738 SerAsnAla----- 2740
Db      9395 GCCAATCAGTGGCAACTGAGAGAGCACACACCTATGAAGTGAAGTGAATCAGATGT 9454
Qy      2740 ----- 2740
Db      9455 CTGGAAGGTTATACAGATGATACAGATTAACATTAATTCCTGTGCAAAAAGATGCTGC 9514
Qy      2740 ----- 2740
Db      9515 TGGTCCCTGAGAAATCTCTGAGATCTTAATAAATATGCTCTCCCGAATCATACAA 9574
Qy      2741 ---ArgValAlaPheSerAspGlyLeuValPheSerSerIleValYrGluCysArg 2759
Db      9575 CATATCTGTGTATGAGGAGCAATTCAGTGTGAATGGCAAGTTTCTGTGTGATGCA 9634

```


QY 760 ProIleSerAlaGlyLeuTyrglyAaenPheThrAlaGluValAlaArgPheIleSerAspPhe 779
Db 3034 TCTCTTACAGC-----AACTTAATTCATAATAAGCTGATATTGTGCTGACTCC 3084
QY 780 SerMetSerTyrglyGlyPheAenIleThrPheSerGlyTyraAspLeuGlyProCysGlu 799
Db 3085 GCCCTTGCCACAGAGCTTTCCATTAATTAAGAGCAATCGAT----- 3129
QY 800 GluProGluValProAlaTySerIleArgGlyLeuGlnPheGlyValGlyAspPhe 819
Db 3130 -----GCATCATCAGTA----- 3141
QY 820 LeuThrPheSerCysPheProGlyTyraArgLeuGluGlyThrAlaArgIleThrCysLeu 839
Db 3141 ----- 3141
QY 840 GlyGlyArgArgArgLeuTrpSerSerProLeuProArgCysValAlaGlyCysGlyAaen 859
Db 3142 -----TCTTAATAT 3150
QY 860 SerValThrGlyThrGlnGlyThrLeuLeuSerProAaenPheProValAenTyraAaen 879
Db 3151 GACTATACAGATTAATTTGGAGTCTCTCCCTCCGAACTTCCCAATATTAACCCAGT 3210
QY 880 AaenHISGlyCysIleTyraSerIleGlnThrGlnProGlyLeuGlyIleGlnLeuValA 899
Db 3211 AACTGGAGTGCATCTACAGAAATCACTGGGACTCAACCAACAGATTCATTCCTTC 3270
QY 900 ArgAlaPheGlyLeuSerGlu-----GlyAspValLeuValTyra 913
Db 3271 ACAGACTTCACTTGGAGACTATTGGGTCAAGTGTGATATTTTGAAGAAATCAG 3330
QY 914 AspGlyAaenAaenSerAlaArgLeuLeuGlyValPheSerHisSerGluMetCysGly 933
Db 3331 GACGGAGGCTACCAAACTGCGCGCTTGCGGATTTACTGTGCTCA--GTTTGGCT 3387
QY 934 ValThrLeuAaenSerHisSerSerSerLeuThrLeuAaenPheIleThrAaenAlaGluAaen 953
Db 3388 CCTACATCATCTCTCCACAGTAACAGCTGTGGCTAAAGTTAAGAGTGAGCGGCACATC 3447
QY 954 ThrSerLeuGly----- 957
Db 3448 ACGGCAAGGGGTTCTCAGCGTACTGGGAGGATCATCAAGAGCTGTGAGGATTAATCTC 3507
QY 957 ----- 957
Db 3508 ACCACCCACAGGTGCTCATGCGCCCACTACCCGATGCCCTACTACACAGCTCCGAA 3567
QY 958 -----PheGluLeuHisPheSerSerPhe 965
Db 3568 TGTCTACTGGCGGCTCGAAGCCAGTCAATGAGCCCTTTCAGAGCTGGAATTCCAAGACTTC 3627
QY 966 GluLeuIleLeuCysGlyAaenProGlyThrProLysPheGlyTyraLysValHisAaenGlu 985
Db 3628 CACTG-----GAACACCAACCCAGCTGCTCTGTGATTACTTGGCGCGTGTGAT 3678
QY 986 GlyHisPheAlaGlySerSerVal-----Ser 994
Db 3679 GGGCCGACTACAACTCCGACTGATAGATTAATGTGTGGGATAGCAACACTGCTCC 3738
QY 995 PheSerCysAspProGlyTyraSerLeuArgGlySerGluLeuLeuCysLeuSerGly 1014
Db 3739 ATCCGTTCCATTAAGAAGCTGTATTGTAAAACTGAGSACTG-----ATGCAAGG 3792
QY 1015 GluArgArgThrTrp-----AspArgProLeuProThrCysValAlaGluCysGlyGly 1032
Db 3793 CAGCT-AGCGCGGTGGCTTGAATCAA--CTTCGCGCAGAGATGAGCAATGTGTGAT 3848
QY 1033 -ThrValArgGlyGluValSerGlyGluValLeuSerProGlyTyraProAlaProTyraG 1052
Db 3849 AGTG-----ACAAAAACCTTTGGCATCTCTGAGAGCATTAATTAATTCCAATCCATATGA 3902

QY 1052 uHisAenLeuAenCysIleTrpThrIleGluAlaGluValGlyCysThrIleGlyLeuH 1072
Db 3903 TAAAGAACCAAGCTTTGTAAGTGAATGAGATCCAAACCAACCCGCAACCCGTGAACATCAAC 3962
QY 1072 sPheLeuValPheAspThrGluGluVal-----HisAspValLeuArgIleTr 1088
Db 3963 GTTTCGTGGATTTGATGTGAAGATTATACATGAATGCTCCACAGATTATGTAGAGCTCTA 4022
QY 1088 PheArgGlyProValGlySerGlyValLeuLeuValLeuLeuSerGlyProAlaLeuProLy 1108
Db 4023 TATATGACCAATGATGAGGGA-----CGCTACTGTGAAATTAATGATCCGCC 4070
QY 1108 sAspLeuHisSerThrPheAaenSerValValLeuGlnPheSerThrAspPhePhe 1128
Db 4071 ACCAGGGGCTACAAACAGCTCCCACTCCAGCTACTGTTCCATACAGATGGATCATATTC 4130
QY 1128 r---LysGlnGlyPheAlaIleGlnPheSerVal-----Se 1139
Db 4131 TGGGAAAAAGATTAAAGATGACAGGTTCACATGAGCTGTGTGGAGAGATGCTGTG 4190
QY 1139 rThrAlaThrSerCysAaenAspProGlyIleProGlnAaenGlySerArgSerGlyAspSe 1159
Db 4191 AACCGAGGCTCTTCAGCAGCCCTGGGTACCCCAACACTATCTCACAAACAAAGAGTG 4250
QY 1159 r---TrpGluAlaGlyAspSerThrValPheGlnCysAspProGlyTyraAlaLeuGln-- 1177
Db 4251 TATCTGAAAC-----ATTGCGGTGGCCCAAGGAGATGATTCAGCT 4292
QY 1178 -----GlySerAlaGluI 1182
Db 4293 CACCATTCATGACTTTGATGTGAATATCATCAAGCTCAACTGATGATCTCCCTGGAGAT 4352
QY 1182 e-----SerCysValLys----- 1186
Db 4353 CTATGAGGTCTTGAATTTTAAGTCTCCCAAGAAATGCCCACTGTGTTCCCAATCACCGTC 4412
QY 1186 ----- 1186
Db 4413 ACGGAACCCCATGACAGTCTCCAGACTGGCAATGAATCAATCCGATTTAAGACGA 4472
QY 1187 -----IleGluAaenArgPhePhe-----TrpGlnProSerProProThrCysI 1201
Db 4473 TAGCACTTAATAATGGAAGAGCTTTCAATGCTCGTGGCAGACAGTCCCT- 4521
QY 1201 eAlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeuSerProAsnTyra 1221
Db 4522 -GAGGTTGTGTGAATTAATCACTTCCAGAGAGAGATTCATTTCCAAATTAATCC 4580
QY 1221 oGluProTyraProProGlyLysGlyCysAspTrpLysValThrValSerProAspTyra 1241
Db 4581 CAACAACATACAGACTAACACAGAGTGTCTCTGATCATTTCAAGTTGACGACATCAACG 4640
QY 1241 IleAlaLeuValPheAaenIlePheAaenLeuGlnProGlyTyraAspPheLeuHisIleTy 1261
Db 4641 TGTCTCTGATATATCACTGACTTGAACCTTCAGATTCCTGCTTCACTGACTTAT 4700
QY 1261 rAspGlyArgAspSerLeuSerProLeuIleLysSerPheTyrgly---SerGlnLeu 1280
Db 4701 GATGCTCAAGTTCCAAACAGCCGCTGCTCCAGTGTGTGGAAGACAGACGCCCC 4760
QY 1280 oGlyArgIleGlySerSerSerSerSerLeuPheLeuAlaPheArgSerAspAlaSerVa 1300
Db 4761 TAACCTATATCATCGCTTCAGAAACAGCTCTTTGTGAGATTCGGTCTGATCTTCCAG 4820
QY 1300 IserAaenAlaGlyPheValIleAspTyraThrGluAaenProArgGlySerCysPheAsp 1320
Db 4821 CCAGAAACAGGGCTTCGGGCTGAATTC-----ACGGAAGAGTGC----- 4860
QY 1320 oGlySerIleLysAaenGlyThrArgValGlySerAspLeuValLeuGlySerSerValTh 1340
Db 4861 -----GAGAGCGCATCATGACGAC-----TCTTCGATATAC 4892
QY 1340 r-----TyraTyraCysHisGlyGlyTyraGluValGluGlyThrSer----- 1353

```
Db 4893 TATCTTCTCCAGCTGTAACCTGACAACTATCTACACAGCAAGAACTGTTCTGGATAT 4952
Qy 1354 -----Thleuser----- 1356
Db 4953 TGAAGCTCAGCCTCATTCATCACTATCTCTCTTTCATGCACTTCAACTTCAAAA 5012
Qy 1357 -----Cys-----IleuGlyProAspGlyys 1364
Db 5013 CAGCAGACTGTACAGCGGACTTTGTAGAAATTTTG-----GATGGCAACAGACTATGA 5066
Qy 1365 -----ProVal-----TyrAsnAspProArgProVal 1373
Db 5067 CGACCTGTCCAGAGCGCTTACTGTGGTTCTCTCCCTGCCCAACCCATCATATGATTTGG 5126
Qy 1373 ----- 1373
Db 5127 CAATGCCCTTACCGGTAGGTTTGTCACTGATTCACAGCAGTTTGGAGGTTCCGTGC 5186
Qy 1374 -----CysThrAlaProCysGlyGlyGlnThrValGlySerAspGlyVal 1389
Db 5187 CATTTATTTGCATGACATGATATCTGTGTGGAAGCTTCTACACACTTGTATGGCATCTT 5246
Qy 1389 IleuSerProAspTyrProGlnAsnTyrThrSerGlyGlnIleCysLeuTyr---PheVal 1408
Db 5247 CAATAGCCCGACTACCCAGACAGACTACCAATCCAAATGCAGATGTGTCTGGAACATTC 5306
Qy 1408 1ThrValProLysAspTyrValValPheGlyGlnPheAlaPhe---PheHisThrAla 1427
Db 5307 CAGCTCCCTGGCAACCGCTGCAACTGCTCTCTTCTTCAATTTGGAGAAATCTCT 5366
Qy 1427 uAsn-----AspValValGlyValHisAspGlyHisSerGlnHisSerArgLeu 1444
Db 5367 AAACTGTAAACAGATTTTGTGAAATCCGAAAGAAATGCC-----AC 5411
Qy 1444 userSerLeuSerGlySerHisThrGlyGlySerLeuPro-----LeuAlaThr 1460
Db 5412 GGGCAGCTTGATGTGAGATCTGAGTGGAACTCCCTCCCTGGGAATTAATTCGTACGTGA 5471
Qy 1460 rSerAsnGlnValLeuIleLysPheSerAlaLysGlyLeuAlaProAlaArgGlyPhe 1479
Db 5472 GGGACATAGTCTATGAGTCCGATTTGTCTGATGCTCAGGACATGCGATGGGCTTCCA 5531
Qy 1479 ----- 1479
Db 5532 GGCAGGTTCAAAATATATTGTCGCAATATATATTGTGGAACTCATGGGAAATCGC 5591
Qy 1479 ----- 1479
Db 5592 ATCTCCCTTCTGGCCTGGAATAACCTTACAACTCCAAATTCAAAATGGGTGAATGT 5651
Qy 1480 -----HisPheValTyrGlnAlaVal-----ProArgThrSe 1490
Db 5652 GAGCAGCATATCATTTATCCACGCTAGATCTTGAAGATGAGCATAGAACCCACAGAA 5711
Qy 1490 r----- 1491
Db 5712 CTGCTTTTATGACAGTTAAAGATTATGATGATTGACACTCATTCCTCCCTCATTTGG 5771
Qy 1491 aThrGlnCysSerSerValProGluProArgTyrGlyLysArgLeuGlySerAspPhe 1511
Db 5772 CACTTACTGT-----GATCCCAAGACAGAAATCTTTTACTC 5807
Qy 1511 rValGlyAlaIleValArgPheGluCysAsnSerGlyTyrAlaLeuGlnGlySerProG 1531
Db 5808 CAGTAGAAACTATCTGACATTCAGATTTCTTCGACACTCTGTGTGACAGAAAG----- 5862
Qy 1531 uIleGlnCysLeuProValProGlyAlaLeuAlaGlnTyr-----AsnVal 1546
Db 5863 -----GATTCCTTCTGGAGTGGTTTGCAGTAGATGTTTCTGA 5900
Qy 1546 1SerAlaProThrCysValValPro-----CysGlyGlyAsnLeuThr-----GluArg 1562
Db 1562 ----- 1562
```

```

QY 1877 -ValLeuThrCysGlnHnIsgLYThrAsnArgAsn-----TrpAs 1889
Db 6634 CTACTTGAATTGGGTGACGAGGCCAACTCGAACGTGTGTTCCAAAGTTGTGTGG 6893
QY 1889 pHISProLeuProLYsCysGluVal----- 1897
Db 6894 CCACACTGTGCTCATAGTGGTATCTCTGAGAGAAAGAAATATATTGAAGTTTCACAC 6953
QY 1898 -----ProLYsGly 1901
Db 6954 TGACGGTGTTCACGCTTACATGAGATTCAGGCCAAGTACTTATAGCTTCTGTGGAGG 7013
QY 1901 YAsnIleThrsSerSerAsnGlyThrValTYrSerProGlyPheProSer---ProTYrSe 1920
Db 7014 AACAGTCTCAGGGGACAGTGGAGTCATCAGAGACATTTGGTACCCGACCTTCCGTTATGC 7073
QY 1920 rSerSerGlnAspCysValTYrLeuIleThrValProIleGlyHnIsgLYValArgLeuAs 1940
Db 7074 AAACAAATGTGTTTGTGATGGTTTATCCGAGGCTCCAGGACACTACCTCACTCTCAG 7133
QY 1940 nLeuSer-----LeuLeuGlnThrGluPro-----SerGlyAspPheIleThrIleTr 1956
Db 7134 TTTTGAAGATTTTAACTTCAAGCTCTCTCGTTGTACAAAGACTTGTGAGATCTG 7193
QY 1956 PAAspGlyProGlnGlnThrAlaProArg---LeuGlyValPheThrsrSerMetAlaTy 1975
Db 7194 GGAA-----AACCATACCTCTGGAGAGATTTCTGGGAGATATTGTGAAATCCCACTCC 7247
QY 1975 sLYsThrValGlnSerSerSerAsnGlnValLeuLeuLYsPheHnIsgLYsPheAlaIaTh 1995
Db 7248 TAGCAGTGTGACACTTCCAGCAATGTGCTCTCGTCAAGTTGTCCACAAAT----- 7299
QY 1995 rGlyGlyIlePheHnIleAlaPheSerAlaTYrProLeuThrLYsCysProProTh 2015
Db 7300 -GGCTCTGTCACTGCCCTCAGGATTTAGG----- 7326
QY 2015 rIleLeuProAsnIleGluValValThrGluAsnGluPheAsnIleGlyAspIleVa 2035
Db 7326 ----- 7326
QY 2035 lArgTYrArgCysLeuProGlyPheThrLeuValGlyAsnGluLeuThrCysLYsLe 2055
Db 7326 ----- 7326
QY 2055 uGlyThrTYrLeuGlnPheGlnGlyGlyProProIleCysE-----GluValHnIscYp 2073
Db 7327 -----CTGAGTTTAAAGTCCAGCAGACAAAGTGTGTGGGGATTTTACATGGCCC 7376
QY 2073 cThrAsnGluLeuLeuThraspSerThrGlyValIleLeuSerGlnSerTYrProGlySe 2093
Db 7377 T-----ACTGGCACTTACTTCTCCCAACTACCCAAACCC 7412
QY 2093 rTYrProGlnPheGlnThrCysSerTYrLeuValArgValGluProAspTYrAsnIleSe 2113
Db 7413 AAATCTCTCATGCCGGATCTGTAGTAGCAGATCACTGACAGAAAGAAAGGCGATCTGT 7472
QY 2113 rLeuThrValGluTYrPheLeuSerGluLYsGln-----TYrAspGluPheG 2129
Db 7473 CCGAGAGTTTACCAACTTGAGGCTGAGTACCCAGCATCTTGAACAGTAGAGACTCAT 7532
QY 2129 uIlePheAspGlyProSerGlyGlnSerProLeuLeuLYsAlaLeuSerGlyAsnTYrSe 2149
Db 7533 CGTATTCATGAGCATTAAGAACAACTGCCCTACTACAGAAACTGTGACGCCGTGTAA 7592
QY 2149 rAlaProLeuIleValThrsSerSerAsnSerValTYrLeuArgTYrPheSerAspH 2169
Db 7593 TGTGACCAATGAATTCAAATCTTCAGAGAAACACCAAGAGTGTATTTTTCATGTATG 7652
QY 2169 sAlaTYrAsnArgLYsGlyPheLYsIleArgTYrSerAlaProTYrCysSerLeuProAr 2189
Db 7653 CTCGGCGCGGTATGAGAGGCTTC-----ACTGCTCTCTACACCTCTACTGAA 7700

```

```

QY 2189 sAlaProLeuHnIsgLYPheIleLeuGlyGlnThrsSerThrGlnProGlyGlySerIleH 2209
Db 7701 TGCAGTGTGTGGATTTCTT-----CCAACTGTCTCCGGTGGAACTTT-- 7746
QY 2209 sPheGlyCysAsnAlaGlyTYrArgLeuValGlyHnIsgLYsMetAlaIleCysThrArgH 2229
Db 7747 -----TCTCTCCGCTATAT-- 7764
QY 2229 sProGlnGlyTYrHnIsgLYsLeuThrPheSerGluAlaIleProLeuCysGlnAlaLeuSerCysG 2249
Db 7765 -----GG 7766
QY 2249 YLeuProGluAlaProLYsAsnGlyMetValPheGlyLYsGlyTYrThrValGlyTYrLy 2269
Db 7767 AATCGGTATTTATGCCAGAAAC-----CTAAGCTGTGAATGAGCTCTCAGTAATCC 7817
QY 2269 sAlaValTYrSerCysSerGlyLY---TYrHnIsgLYsAlaGlyAlaIleThrAl 2288
Db 7818 AATCGGGAATAATTCATCATAGTATCTATTTCTAGAACTTCCATTGAAGATCATCA 7877
QY 2288 sGlyCysLeuAspThrGlyLeu----- 2295
Db 7878 AGACTGTACATTGATATGCTCTTGAAGTTTCAGATGAGGAGATGAGGCCCTGATAGA 7937
QY 2296 ---TrpSerAsnArgAsnValProProGlnCysValProValThrCysProAsp----- 2312
Db 7938 GAAGTTCTGTACCTGTGACAGCAACAGCGCCCTTGATGATCCCTACCTCAGAGTGG 7997
QY 2313 -----ValSerSerIleSerValGluHnIsgLYsIleArgTYrArgLeuIlePheGluTh 2329
Db 7998 GATACGCTTCGTACAGCAATGAGCGGTAGAAATATCTGGA-----TTCTATAT 8045
QY 2329 rGlnTYrGlnPheGlnAlaGlnLeuMetLeuIleCysAspProGlyTYrTYrTYrThG 2349
Db 8046 AGAGTACTCTT-----ACAATTTGTGTGAATTCGAGACAG 8084
QY 2349 YGlnArgValIleArgCysGlnAlaAsn-----GlyLYsTrpSerLeuG 2364
Db 8085 TCACATAGAGATGATCTCAAGTCTTATTCATTCACAACTTGTACAGTGCATGAGACC----- 8139
QY 2364 YAspSerThrProThrCysArgIleIleSerCysGlyGluLeuProIleProProAsnG 2384
Db 8140 -----CACTGTTCAATGCTGCTGAAAGCCCGAAGG 8171
QY 2384 YHnIsgLYIleGly---ThrLeuSerValTYrGlyAlaThrAlaIlePheSerCysAsnSe 2403
Db 8172 GCACACATCACTCACTCACTCAGTACTTCTTCCAGGCTCATCCAACTTGACATTC 8231
QY 2403 rGlyTYrThrLeuValGlySerArgValArgGlyCysMetAlaAsnGlyLeuThrPheSerG 2423
Db 8232 AGACTCCGCTCACTGTC-----AGAAATGTGTACTCCCCAG 8267
QY 2423 YSerGlyValArgCysLeuAlaGlyHnIsgLY---ThrProGluProIleValAs 2441
Db 8268 ATCGCCGCTC-----ATAGAGCATGACTGTGAGAGTCACTCCAGCGGATACAGTGC 8321
QY 2441 nGlyHnIsgLYs-----IleAsnGlyGluAsn----- 2448
Db 8322 TGGTTCCAACCAACTTATAGTACTTTTAAACAACAATCAAGGGCAAACCTCGGAT 8381
QY 2449 -TYrSerTYrArgLYsSerValValTYrGlnCysAsnAlaGlyPheArgLeuIleGlyMe 2468
Db 8382 TTATTCACAATGAGACCAACAAGCGTTTGTGGGGGAACATTCCTCAGCTAATAG 8441
QY 2468 tSerValArg----- 2471
Db 8442 TACATCAAAATCTCTCACTGCGCTCAGACATTCAGAAAAACAGCAGATGCTCTGGAC 8501
QY 2472 ---IleCysGlnGlnAspHnIsgLYsTrp-----SerGlyLYsThrProPheCysValPr 2488
Db 8502 AGTATCACTACCATTAAGTAAACACTGGAGATTAAGCTTTGACAGCAATTTCCGAATCCC 8561
QY 2488 O----- 2488

```

Db 8562 CAGCAGTGAAGCAGAGTGTGAGAAAGCTTGTGAAGTTGGGAGGAGCTTGATGAT 8621
 Qy 2489 -----Ilethrcysgilyhisprogly---AsnProValaAsnGlyLe 2501
 Db 8622 CAATAGAACCTGTGTAGCAGAGAGCTGTGAGATGGCTCCAGTCCCAT-----GT 8675
 Qy 2501 uThrGlnGlyAsnGlnPheAsnLeuAsnAspValValLysPheValCysAsnProGlyTy 2521
 Db 8676 CACATCAGGGAGAACATTTTC-----ACTGCTGTCT 8705
 Qy 2521 rMetAlaGlnGlyAlaAlaArgSerGlnCysLeuAlaSerGlyGlnTrpSerAspMetLe 2541
 Db 8706 CCAATCTGAGAGATGGCA-----GCCAGGCTTCTGTGATCTTCAT 8750
 Qy 2541 uProthrcys---ArgIleIleAsnCysThrAspProGly---HisGln----- 2555
 Db 8751 TAGCCGGTGGCGAGACATTCAT---ACCTCCCTGTGTGACATCATCTCCAAACT 8807
 Qy 2556 GluAsnSerValArgGlnValAlaAlaSerGly---ProHisArg----- 2569
 Db 8808 CCCGAGCAATAGCAGAACAAACATGAATGCACTTACCTATAGACGCTGACCTGAGTC 8867
 Qy 2570 -----PheSerPheGlyThrThrValSerTyArgCysAsn 2581
 Db 8868 TCTGTGATCCTGACTTTTGTGCTCTTCATTTGGAAGATGCTCAGCTACACCGGAGC 8927
 Qy 2582 -----HisGlyPheTyLeuLeuGlyThrProValLeuSerCysGlnGly 2596
 Db 8928 CTGTGATCATGATGCTGTGCACATCATCAAGAGTCGTAACCTCTTCCATCCTCTGCT 8987
 Qy 2597 AspGlyThrTrp-----AspArgProArgProGlnCysLeuLeuValSerCysGlyHis 2614
 Db 8988 GACCATATGTGTGTTCTGAACCTTGGCTCCCTCACTGT----- 9026
 Qy 2615 ProGlySerProProHisSerGlnMet-SerGlyAspSerTyThrVal-----GlyAl 2632
 Db 9027 ---GGAGGGCCCACTGTTGCTCAACTTCTATCTGATGATGATGATGATGATGATGATG 9083
 Qy 2632 aValValArgTySerCysIleGlyLysArgThrLeuValAlaLeuSerThrArgMetCys 2652
 Db 9084 CAAGATTTCCATAGAGAGCATCACCCTGTGTGAGATCAATGAATCTCTGGAAATCT 9143
 Qy 2652 sGlyLeuAspGlyHisTrpThrGlySerLeuProHis----- 2664
 Db 9144 TAGAGCCCTTCTACTCATAGACAACTACCCCAACAACTCTACTGTGTCTAGACCT 9203
 Qy 2664 ----- 2664
 Db 9204 CCATGTTAGAGCAGCAGAGATGAATTAATTAAGTTCAATGATTCGATGCTCCTTC 9263
 Qy 2665 ---CysSerGlyThrSerValGlyValCysGlyAspProGlyIleProAlaHisGly 2682
 Db 9264 CAACCTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9323
 Qy 2682 yIle---ArgLeuGlyAspSerPheAspProGlyThrValMetArgPheSerCysGlyAl 2701
 Db 9324 TCTTGAAAGTCTGTGTGTTCCAGCGCTCAAACTGTTAAGACCAATAGACAGCT 9383
 Qy 2701 aGlyHisValLeuArgGlySerSerGlyArgThrCysGlnAlaAsnGlySerTrpSer- 2720
 Db 9384 GACCTGTGTTCAAGACAGATTTCTTCAACAGCAAGA-----GGTTGGAAAT 9444
 Qy 2721 -----GlySerGlnProGlnCysGlyValIle----- 2729
 Db 9435 ATTTTCCGAGAGCAATAGGGCCACAGAGAGATGTGTGATACCTGACGAGAGCAA 9494
 Qy 2730 ---SerCysGlyAsnProGlyThrProSerAsnAlaArgVal----- 2742
 Db 9495 CCAGAGCTTTGTGTCTCTGATTTGATTTGAATGAGGCTATGACAAAGGCTTCAGCTG 9554
 Qy 2743 -----ValPheSe 2745

Db 9555 CATATGATACATAGTTGACCTGAAAAAACAACCTGTTAAGCTCACCTTCAATGTGTTCA 9614
 Qy 2745 r---AspGlyLeuValPheSerSerIleValTyrgLysArgGlyGlyTyTyTyAl 2764
 Db 9615 TCTGAGAGGACCATTCGTGACCTGGAGCTCGCTTATGATTAATGATGAG----- 9663
 Qy 2764 aThrGlyLeuLeuSerArgHisCysSerValAsn-----GlyThrTrpThrGly 2780
 Db 9664 -----ATAGCAATGGCCCAAGCAATTAATCTATTTAGTGGAAAAATTCGTGTG 9713
 Qy 2780 ySerAspProGlnCysLeuValIle----- 2788
 Db 9714 CTCCCGTAT-GCCTGCCCCATTTATCTCTCCGCTACTTCTTACGTTTCAGTTGTCT 9772
 Qy 2789 ---AsnCysGlyAspProGlyIleProAlaAsnGlyLeuArgLeuGlyAsnAspPh 2806
 Db 9773 CTGACGTAACTGTGAATAGAGGATTAATGCAATATACCTTTGTGAGACATGCTT 9832
 Qy 2806 eArgTyAsn-----LysThrValThrTyrgLysCysValProGlyTyTyMetMetGly 2823
 Db 9833 GCGGGGGAACATATAGCAACCACTCGACACTTCAMAAATGCGTATCATCTCATTTATCA 9892
 Qy 2823 uSerHisArgValSerValLeu-SerCysThr----- 2833
 Db 9893 A---CATCGAGCGGCATATCTCCACTGTACTTGGGTGATGCGAGCTCCCCACAGAGC 9949
 Qy 2834 ---LysAspArgThrTrpAsnGlyThrTyTyProVal-----CysValAlaLeuMetC 2850
 Db 9950 AGGTTCAATAGTATGTTGTGAGCTTACAGCTGCTTCACAAAGCTGTCTCAAAAGTACT 10009
 Qy 2850 yLysProProProLeuIleProAsnGlyLysValValGlySerAspPheMetTrpGly 2870
 Db 10010 TAGAATCTCAGATTCAGTACAGACTGT-----GGA 10042
 Qy 2870 eSerValThrTyAlaCysLeuGlnGlyTyrgLysLeuSerLeuProAlaValPheThrC 2890
 Db 10043 ACCGGGTGACTAGTTCTGGGTGCGCAATAT---ACAACCTTGGCAGTGTTCATCC- 10098
 Qy 2890 yGlnGlyAsnGlySerTrpThrGlyGlnLeuProGlnCysPheProValPheCysGly 2910
 Db 10099 -----TCAAATGAGCACTGGGTGCTGTTTC-----A 10126
 Qy 2910 sProGlyValProSerArgGlyArgArgGlyAspArgGlyPheSerTyArgSerSerV 2930
 Db 10127 AGTCTGAGATTATTAACAGAACTCGCAAGTCAA-----TCTCTTATCAAGTTGCA- 10179
 Qy 2930 aSerPheSerCysHisProProLeuValLeuValGlySerProArgArgPheCysGln 2950
 Db 10180 -----GATTCGAAC-----AGAGATPACAAACCAA 10204
 Qy 2950 eAspGlyThrTrpSerGlyThrGlnProSerCysIleAspProThrLeuThrTyCysA 2970
 Db 10205 CGTTTGGC-----AATCTGA 10219
 Qy 2970 IaAspProGlyValProGlnPheGlyTyrIleGlnAsnAsnSerGlnGlyTyrgValGly 2990
 Db 10220 AGACTCTGTGTGCTCAG---AATATATACAAATAC-----CTGACT 10261
 Qy 2990 eThrValLeuPheArgCysGlnGly-----TyTyLeuLeuG 3003
 Db 10262 GCAACATCATTCACAGAGCCCGAGAACACAGACATTCCTCTTTTCTATTTGCTTTC 10321
 Qy 3003 lngLysThrThrArgThrCysLeuProAsn-LeuThrTrpSerGlyThrProProAsp 3022
 Db 10322 AGCTGGAAGATTCAGACAAATGATG---AATATTTCTTGAGAGTAAGAAAGCGCGCA 10378
 Qy 3023 CysValProHisHisCysArgGlnProGlyThr-----ProThrHisAlaAsn 3038
 Db 10379 GCAGACCTTACCACTGCTTGAACAAGTACTGTAGCAACTGTGCCCCAGCGCTTCT 10438
 Qy 3039 ValGlyAlaLeuAspLeu-----ProSerMetGly 3048
 Db 10439 CTCAGAGCAACAACTGTATCTGCACTTTCAGAGCGACCACTGATCACCACAAATGGCT 10498

Db	451	TCGGCAACATGCATCTCAGAGTACTGTCACTTATGGGATTAATGAACACCTTATTTGT	510
Qy	854	--ValAlaGluCysGlyAsnSerValThrGlyThrGlnGlyThrLeuSerProAsn	872
Db	511	GACAGAAATCCTTGTGGGCTACCCCCACCATCACAATGGAGATTTCATTGACACCAAC	570
Qy	873	PheProValAsnTyrAsnAsnAsnHisGluCysIleTyrSerIleGlnThrGlnProGly	892
Db	571	-----AGAGAAATTTTCACATGAGATCAGCGTACGCTACCGCTGCATCTCTGGA	621
Qy	893	LysGlyIleGlnLeuValAlaArgAlaPheGluLeuSerGluGlyAspValLeuLysVal	912
Db	622	ACCGGA-----GGAGAAAGGTGTTTGACTT---GTGGGTGAGGCC---TCCATA	666
Qy	913	TyrAspGlyAsnAsnAsnSerAlaArgLeuGluGlyValPheSerHisSerGluMetMet	932
Db	667	TACTGCACACGCAATGACGATCAA-----GTGGCATCTGGAGCGGCCGCCCTCAG	720
Qy	933	GlyValIThrLeuAsnSerThrSerSerSerLeuThrPheAspPheIleThrAspAlaGlu	952
Db	721	TGCATTATACCTTAAC-----	735
Qy	953	AsnThrSerLysGlyPheGluLeuHisPheSerSerPheGluLeuIleLysCysGluAsp	972
Db	736	-----AAATGCAGCCT	747
Qy	973	ProGlyThrProLysPheGlyTyrLysValHisAspGluGlyHisPhe-----AlaGly	990
Db	748	CCAAATGTGAAAAAT---GGAATATGTGTATCTGCACACAGAGCTTATTTCTTAAAT	804
Qy	991	SerSerValSerPheSerCysAspProGlyTyrSerLeuAsnGlySerGluLeuLeu	1011
Db	805	GAACTGTGGAGCTTTAGGTGTCAAGCTGGCTTTGTCATGAAGAAGCCCCCGCTGTGAG	864
Qy	1011	CysLeuSerGlyGluAsnArgThrThrAspArgProLeuAsnProThrCysValAlaGluCys	1030
Db	865	TGCCAGGCGCCGAAACAA-----TGGGAGCGGAGAGCTACCAAGCTGCTCCAGGTTAGT	918
Qy	1031	GlyGlyThrValArgGlyGluValSerGlyGlnValLeuSerProGlyTyrProAlaPro	1050
Db	919	-----CACCCACCTCCA	930
Qy	1051	TyrGluHisAsnLeuAsnCysIleTyrThrIleGluAlaGluAlaGlyCysThrIleGly	1070
Db	930	-----	930
Qy	1071	LeuHisPheLeuValPheAspThrGluGluValHisAspValLeuArgIleTyrAspGly	1090
Db	931	-----GATCGTCGCAAT-	942
Qy	1091	ProValGluSerGlyValLeuLeuLysGluLeuSerGlyProAlaLeuProLysAspLeu	1111
Db	942	-----	942
Qy	1111	HisSerThrPheAsnSerValIleGlnPheSerThrAspPhePheThrSerLysGln	1130
Db	942	-----	942
Qy	1131	GlyPheAlaIleGlnPheSerValSerThrAlaThrSerCysAsnAspProGlyIlePro	1150
Db	943	-----GCT	945
Qy	1151	GlnAsnGlySerArgSerGlyAspSerTyrGluAlaGlyAspSerThrValPheGlnCys	1170
Db	946	GAGGAGTACCAAAAGGAGCAAGACAACTTTTACCTTGAGGAGAAAGTGTCTTACAGCGT	1005
Qy	1171	AspProGlyTyrAlaLeuGlnGlySerAlaGluIleSerCysValLysIleGluAsnArg	1190
Db	1006	GAGCCCGGCTACACCTCAGAGGGGCTGCTTATGCGCTGCACACCCAGGAGAC---	1063
Qy	1191	PhePheTyrGlnProSerProProThrCysIleAlaProCysGlyGlyAspLeuThrGly	1210

Db 1063 -----TGAGCCCTGAGACCCCAACATGTGAAGTAAATCCTGTGATGATCATGGGC 1116
 QY 1211 Pro-----SerGlyValIleLeuSerProAsnTyrProGluProTyrProProGlyLys 1228
 Db 1117 CAACCTTCTTAATGCGCGGTGCTATTTCCTCA----- 1146
 QY 1229 GlucyAspTrpLysValThrValSerProAspTyrValIleAlaLeuValPheAsnIle 1248
 Db 1146 ----- 1146
 QY 1249 PheAsnLeuGluProGlyTyrAspPheLeuHisIleTyrAspGlyArgAspSerLeuSer 1268
 Db 1147 GTAAATCTCCACGCTTGAGCAAAAGTGAGATTGTGTGAT----- 1188
 QY 1269 ProLeuIleGlySerPheTyrGlySerGlnLeuProGlyArgIleGluSerSerSerAsn 1288
 Db 1189 -----GAGGATTTCAATTAAGGC-----AGCTGTGCTAAGT 1221
 QY 1289 SerLeuPheLeuAla-----PheArgSerAspAlaSerValSerAsnAla 1303
 Db 1222 TACTGTGCTTGGCTGGAAATGAAAACCTTTGGAATGACAGTCTCCAGTGTGAAACA 1281
 QY 1304 GlyPheValIleAspTyrThrGluAsnProArgGluSerCysPheAspProGlySerIle 1323
 Db 1282 ATCTTT-----TGTCAGAGCTCCAGTTATT 1308
 QY 1324 LysAsnGlyThrArgValGlySerAspLeuLysLeu-----GlySerSerValThr 1340
 Db 1309 CCTAATGAGGAGACACACAGGAAACCTCTGGAAGTCTTCCCTTGGAAAAGCAGTAAT 1368
 QY 1341 TyrTyrCysHis-----GlyGlyTyrGluValIleGluIleThrSerThr 1354
 Db 1369 TACCATGCGACCCCAACCCAGACAGAGGAGCGAGCTTCGACCTCATTTGGAGAGACACC 1428
 QY 1355 LeuSerCysIleLeuGlyProAspGlyLysProValIlePheAsnProArgProValCys 1374
 Db 1429 ATCGGCTGCACAAAGTGAACCTCAAGGGAATGGGGTTGGAGCACCCGCGCTCGTGT 1488
 QY 1375 ThrIleProCysGlyGlyGlnTyrValGlySerAspGlyValIleLeuSerProAsnTyr 1394
 Db 1489 -----GGAAATTCG----- 1497
 QY 1395 ProGlnAsnTyrThrSerGlyGlnIleCysLeuTyrPheValThrValProLysAspTyr 1414
 Db 1497 ----- 1497
 QY 1415 ValValPheGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspValIleGluValHis 1434
 Db 1497 ----- 1497
 QY 1435 AspGlyHisSerGln-----HisSerArgLeuLeuSerSerLeuSerGlySerHisThr 1452
 Db 1498 ---GGTCACTGTCAAGCCCAAGATCATTTTCTGTTGGCAAGTGAACCAACCAACCAAT 1554
 QY 1453 GlyLeuSerLeuProLeuAlaThrSerAsnGlnValLeuIleLysPheSerAlaLysGly 1472
 Db 1555 GCATCTGACTTTCCCATTTGGGACATCT-----TTAAAGTACGAATGCCGCTCT 1602
 QY 1473 LeuAlaProAlaArgGlyPhe-----HisPheValTyrGlnAlaVal 1486
 Db 1603 GAGTACACGAGGAGGACATTTCTATCATCATGTCTAGTAACTCGTCTGTCTCAAGT--- 1659
 QY 1487 ProArg-----ThrSerAlaThrGlnCysSerSerValProGluProArgTyrGlyLys 1504
 Db 1660 CCCAAAGATCTCTAAACGTAAATCATGTAAACTCTCCAGATCCAGTAAATGGCATG 1719
 QY 1505 ---ArgLeuGlySerAspPheSerValGlyValIleValArgPheGluCysAsnSerGly 1523
 Db 1720 GTGATGTGATCATCAACATCCAGGTGGATCCAGATCAACTATTTCTTGACTACAGGG 1779
 QY 1524 TyrAlaLeuGlnGlySerProGluIleGluCysLeuProValProGlyAlaLeuAlaGln 1543
 Db 1780 CACGAGCTCATTTGGTCACTCATCTGCTGAATGTATC---CTCTGGGGGAATGCGCCCAT 1836

QY 1544 TrpAsnValSerAlaProThrCysVal-----ValProCysGlyGlyAsnLeuThrGluArg 1562
 Db 1837 TGGAGCAGCAAGCCGCAATTTGTCAAGAAATCTTTGGGCTACACCCCAACATGGCC 1886
 QY 1563 ArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeuAsnCysVal 1582
 Db 1897 AATGAGATTTCAATTAGCACCAACAGAGAGAAATTTTCATATGGATCA----- 1944
 QY 1583 TrpLysIleValValProGluGlyAlaGlyIleGlnIleGlnValIleSerPheValThr 1602
 Db 1945 -----GTGCTGACTACCTCCGCTGC 1962
 QY 1603 GluGlnAsnTrpAspSerLeuGluValPheAspGlyAlaAspAsnThrValThrMetLeu 1622
 Db 1963 AATCTCGAAACCGAGGAGAGAAAGTGTTCAG-----CTTGTC 2001
 QY 1623 GlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerThrSerAsnGlnLeuTyr 1642
 Db 2002 GTTGAGCCCTC-----ATATACGCAACCCAGCATGAC----- 2034
 QY 1643 LeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHisLeuGluTyrLysThr 1662
 Db 2035 -----GATCAA 2040
 QY 1663 ValGlyLeuSerSerCysProGluPro-----AlaValProSer----- 1675
 Db 2041 GTGGGCAATCTGGAGCGCGCCGCTCAGTGCATTAATCAACAAATCAGCGCTCCA 2100
 QY 1676 -----AsnGlyValLysThrGlyGluArg-----TyrLeuValAsnAspVal 1689
 Db 2101 AATGTGAATAATGAAATTTGTATCTGACACAGAAAGCTTATTTCTTAATGAAGTT 2160
 QY 1690 ValSerPheGlnCysGluProGlyTyrAlaLeuGlnGlyHisAlaHisIleSerCysMet 1709
 Db 2161 GTGGAGTTTGGTGTGCTCAGCTGCTTGTATGAAGAGACCCGCGCTGGAAGTGCAG 2220
 QY 1710 ProGlyThrValArgAspTyrAsnTyrProProProLeuSerCysIleAlaGlnCysGlyGly 1729
 Db 2221 ---GCCCTGAACAAATGGAGCGGAGCGGACCTCAAGCTGCTCCAGGGTATGT--- 2268
 QY 1730 ThrValGluGluMetGluGlyValIleLeuSerProGlyPheProGlyAsnTyrProSer 1749
 Db 2268 ----- 2268
 QY 1750 AsnMetAspCysSerTrpLysIleAlaLeuProValGlyPheGlyAlaHisIleGlnPhe 1769
 Db 2268 ----- 2268
 QY 1770 LeuAsnPheSerThrGluProAsnHisAspTyrIleGluIleArgAsnGlyProTyrGlu 1789
 Db 2269 ---CAGCCACCTCCAGAGTCTCTG----- 2289
 QY 1790 ThrSerArgMetMetGlyArgPheSerGlySerGluLeuProSerSerLeuLeuSerThr 1809
 Db 2289 ----- 2289
 QY 1810 SerHisGluThrThrValTyrPheHisSerAspHisSerGlnAsnArgProGlyPheLys 1829
 Db 2290 -----CATGCTGAGGATCCCAAAG----- 2310
 QY 1830 LeuGluTyrGlnAlaTyrGluLeuGlnGluCysProAspProGluProPheAlaAsnGly 1849
 Db 2311 -----GACAGAGCAACTTT----- 2325
 QY 1850 IleValArgGlyAlaGlyTyrAsnValGlyGlnSerValThrPheGluCysLeuProGly 1869
 Db 2326 -----TACCCGGGACAGAAAGTGTTCACAGCTGTGACCCGGC 2364
 QY 1870 TyrGlnLeuThrGlyHisProValLeuThrCysGlnHisGlyTyrAsnArgAsnTyrAsp 1889
 Db 2365 TATGACCTCAGAGGGGCTGCTGATACGCTGCACA-----CCCCAGAGAGACTGAGC 2418

QY 1890 HisProLeuProLysCysGluValProCysGlyGlyAsnIleThrSerSer-----Asn 1907
 DB 2419 COTGAGCCGCCACATGTGAAGTGAATCCTGTATGATGCTTCAAGGCCAATCTTAAAT 2478
 QY 1908 GlyThrValIleYrSerProGlyPheProSerProYrSerSerSerGlnAspCysValTrp 1927
 DB 2479 GGCCTGTGCTAATTTCCA----- 2496
 QY 1928 LeuIleThrValProIleGlyHisGlyValArgLeuAsnLeuSerLeuLeuGlnThrGlu 1947
 DB 2497 ---GTAATCTCCAGCTTGGACCAAAAGT--- 2523
 QY 1948 ProSerGlyAspPheIleThrIleTrpAspGlyProGlnGlnThrAlaProArgLeuGly 1967
 DB 2524 ---GATTTTGTGTGAT---GAAGATTTCATTAAAGGC----- 2559
 QY 1968 ValPheThrArgSerMetAlaLeuValThrValGlnSerSerSerAsnGlnValLeuLeu 1987
 DB 2560 ---AGCTCTGCTAGTTTATTTGTGCTTG 2583
 QY 1988 LysPheHisArgAspAlaIleThrGlyGlyIlePheAlaIleAlaPheSerAlaTrpPro 2007
 DB 2584 ---GCTGAATGGAAAGCCTTGGAAATGACAGAGTTCCA 2619
 QY 2008 LeuThrLys-----CysProProProThrIleLeuProAsnAlaGluValIleValThr 2024
 DB 2620 GTGTGTAAACAAATCTTTTGTCCAAAGCTTCAGTATTCCTAATGGGAACACACAGCA 2679
 QY 2025 GluAsn---GluGluPheAsnIleGlyAspIleValArgTrpArgCysLeuPro----- 2041
 DB 2680 AAACCTGTGAAGCTTTCCTTTGGAAAGACGTAATTAACATCGACCCCAACCCA 2739
 QY 2042 ---GlyPheThrLeuValGlyAsnGluIleLeuThrCysLysLeu----- 2055
 DB 2740 GACAGAGGACGACCTTCGACCTCATTTGGAGAGAGCACCCTCCGCTGCAACAGTACCT 2799
 QY 2056 ---GlyThrTrpLeuGlnPheGluGlyProProProIleCysGluVal----- 2070
 DB 2800 CAAGGGAATGGGGTTTGG-----AGCAGCCCTCCCTCGCTGTGAATTTGGGGT 2850
 QY 2071 HisCysProThrArgAsnIleLeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTrp 2090
 DB 2851 CACTGTCAAGC----- 2862
 QY 2091 ProGlySerTrpProGlnPheGlnThrCysSerTrpLeuValArgValGluProAspTrp 2110
 DB 2863 ---CCAGAT--- 2868
 QY 2111 AsnIleSerLeuThrValGluTrpPheLeuSerGluLysGlnTrpAspGluPheGluIle 2130
 DB 2869 ---CATTTCTGTGGCCAAAGTTGAAACCCAAACCAATGCA 2907
 QY 2131 PheAspGlyProSerGlyGlnSerProLeuLeuValAlaLeuSerGlyValSerAla 2150
 DB 2908 TCTGACTTTCCATTGGACATCTTTAAAGTACGAATCCCGTCTGAGTACACGGGAG 2967
 QY 2151 ProLeuIleValIleHisSerSerSerAsnSerValTrpLeuArgTrpSerSer---AspHis 2169
 DB 2968 CCATTCTCTATCACAATGCTAGATACTGGTCTC-----TGCATCAAGTCCCAAAAGT 3018
 QY 2170 AlaTrpAsnArgGlyGlyPheLysIleArgTrpSerAlaProTrpCysSerLeuProArg 2189
 DB 3019 GTCTGTAAAGTTAAATCA-----TGTAAATCTCTCA 3051
 QY 2190 AlaProLeuHisGlyPheIleLeuGlnIleThrSerThrGlnProGlyLysSerIleHis 2209
 DB 3052 GATCCAGTAAATGCGATGCGATGTGATCACAAGACATCCAGAGTTGATCCAGAAATCAAC 3111
 QY 2210 PheGlyCysAsnAlaGlyTrpArgLeuValGlyHisSerMetAlaIleCysThrArgHis 2229
 DB 3112 TATTTCTTGATCTACAGGAGCAGCATCTGATCATCTGCTGAATGATATCTCTCA 3171
 QY 2230 ProGlnGlyTrpHisLeuThrPheSerGluAlaIleProLeuCysGlnAlaLeuSerCysGly 2249

DB 3172 GGCATATCTGCCAT---TGAAGACGAAGCCGCCAATTTGTCAAGAAATCTCTTGAG 3228
 QY 2250 LeuProGluAlaProLysAsnGlyMetValPheGly-----LysGluTrpThrVal 2266
 DB 3229 CTACCCCAACCATGCGCAATGAGATTCATTAGCACCAAGAGAAATTTCACTAT 3288
 QY 2267 GlyThrLysAlaValIleYrSerCysSerGlyGlyTrpHis-----LeuGlnAla 2282
 DB 3289 GGAATGAGTGAACCTTACCGCTGCAATCTTGAAGCAGAGGAGAAAGGTGTTGAGCTT 3348
 QY 2283 GlyAlaGluAlaThrAlaGluCysLeu-----AspThrGlyLeuThrPheSer 2297
 DB 3349 GTGGGTGAGCCCTCCATATCTGACACAGCAATGACATCAAGTGAAGTGTGAGCGGC 3408
 QY 2298 ---AsnArgAsnValProPro----- 2303
 DB 3409 CCCGCCCTCAGTGCATTATACCTTAACAAATGACAGCCTCCAAATGTGAAATGGAATA 3468
 QY 2304 ---GlnCysVal 2306
 DB 3469 TTGTATCTGACACAGAACTTATTTCTTAAATGAAGTTGTGAGTTAGTGTCTAG 3528
 QY 2307 Pro-----ValThrCys----- 2310
 DB 3529 CTTGGCTTTGTATGAAGAGACCCCGCGGTGTAAGTGCAGGCCCTGAACAATGGCAG 3588
 QY 2311 ProAspValSerSerIleSer-----ValGluHisGlyArg 2322
 DB 3589 CCAGAGTTACCAAGTGTCTCCAGGGGTGTGACGCCGCTCCAGAAATCTGTAGTGTAG 3648
 QY 2323 TrpArgLeuIlePheGluThrGlnTrpGlnPheGlnAlaGlnLeuMetLeuIleCysAsp 2342
 DB 3649 CATPACCCCAAGCATACAGACAACTTTTCACTCGGCGAGAGAGTTTATACAGCTGTAG 3708
 QY 2343 ProGlyTrpTrpTrpGlyGlnArgValIleArgCysGlnAlaAsnGlyLeuTrpSer 2362
 DB 3709 CCTGTGTATGACCTGAGAGGGGTGCGTCTGCACTGACACACCCAGAGACAGTGGAGC 3768
 QY 2363 LeuGlyAspSerThrProThrCysArgIleIleSerCys-----GlyGluLeu 2378
 DB 3769 ---CCTGAAGCCCGAGATGTGCAAGTAACTCTGTGATGACTTCTTGCGTCAACTC 3822
 QY 2379 ProIleProAsnGlyHisArgIleGlyIleThrLeuSerVal---TrpGlyValAlaThrAla 2397
 DB 3823 ---CTCATGGCCGTGTGATTTCACTTAATCTCAAGCTTGGGCGCAAGGTG 3873
 QY 2398 IlePheSerCysAsnSerGlyTrpThrLeuValGlySerArgValArgGluCysMetAla 2417
 DB 3874 TCTTTGTCTGTATGAAGGGTTTGGCTTAAGGCGAGTTCCGTTAGTCACTGTCTCTG 3933
 QY 2418 AsnGly-----LeuTrpSerGlySerGluValArgCysLeuAlaGlyHisCysGly 2434
 DB 3934 GTTGAATGAGAGCGTTTGAATATACAGTCTTCTGTGTGTAACATATCTTTTGTCCA 3993
 QY 2435 ThrProGluProIleValAsnGlyHisIleLeuAsnGlyGlu-----AsnTrpSerTrp 2451
 DB 3994 AATCTCAAGTATCTTAATGAGAGACACAGAACTCCCTGTGAGATATTCCTAT 4053
 QY 2452 ArgGlySerValIleTrpGlnCysAsn-----AlaGlyPheArgLeu 2465
 DB 4054 GGAAGAAATATATCTTACATGATGACCCCAACCCAGACAGAGGATGACCTTCAACTC 4113
 QY 2466 IleGlyMetSerValArgIleCysGlnGlnAspHisHis-----TrpSerGly 2481
 DB 4114 ATTGGAGAGACACATCCGTTGCAAAAGTGAACCTTATGGATGGGGTTTGGAGCAGC 4173
 QY 2482 LysThrProPheCys---ValProIleThrCysGlyHis----- 2493
 DB 4174 CCGGCCCTGCTGTAACTTTCTGTTCTGCTGTGCTGATCACTGTAACCCAGACAGCTTT 4233
 QY 2494 ---ProGlyAsnProValAsnGlyLeuThrGlnGlyAsnGlnPheAsnLeu 2509

; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pc_fl_genes Version 1.0
 ; SEQ ID NO 259
 ; LENGTH: 7313
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (13) .. (6147)
 ; US-09-620-312D-259

Alignment Scores:

Pred. No.:	1,3e-93	Length:	7313
Score:	1371.00	Matches:	538
Percent Similarity:	32.87%	Conservative:	266
Best Local Similarity:	22.00%	Mismatches:	832
Query Match:	8.07%	Indels:	810
DB:	4	Gaps:	109

US-10-016-248-2 (1-3104) x US-09-620-312D-259 (1-7313)

```

QY 798 CysGluGluProGluValProAla-----TyrSerIleArgLysGlyLeuGlnPhe 814
DB 337 TGTGTAATCCTCCAGATCCGTGATGGCATGGTGATGATCAAGGACATCCAGTTC 396
QY 815 GlyValAlaGlyAspThrLeuThrPheSerCysPheProGlyTyrArgLeuGlnGlyThrAla 834
DB 397 -----GGATCCCAATTAATATATTCTGTACTAAGATACCGACTCATTTGGTCTCTG 450
QY 835 ArgIleThrCys---LeuGlyGlyArgArgArgLeuThrPheSerCysProLeuProArgCys 853
DB 451 TCTGCCCATGTCATCATCTCAGTGGATCTGATCTGATTTGGATTAATGAACACCTATTGT 510
QY 854 ---ValAlaGlyCysGlyAsnSerValThrGlyThrGlnGlyThrLeuLeuSerProAsn 872
DB 511 GACAGAAATTCCTTGTGGGGTACCCGCCCATCAGCAATGAGATTTCATTACCAAC 570
QY 873 PheProValAsnTyrAsnAsnAsnHisGlyCysIleTyrSerIleGlnThrGlnProGly 892
DB 571 -----AGAGAAATTTTCACTATGATGATGATGATGATGATGATGATGATGATGAT 621
QY 893 LysGlyTyrLeuLeuLysAlaArgAlaPheGlyLeuSerGlnGlyAspValLeuLysVal 912
DB 622 AGCGGA-----GGAGAAAGGTGTTGACCTT---CTGGGTAGAGCC---TCCATA 666
QY 913 TyrAspGlyAsnAsnAsnSerAlaArgLeuLeuGlyValPheSerHisSerGluMet 932
DB 667 TACTGCACCGACGATGATGATCAAA-----GTGGGATCTGGAGCGGCCGCCCTCAG 720
QY 933 GlyValThrLeuAsnSerThrSerSerSerLeuThrLeuAspPheIleThrAspAlaGlu 952
DB 721 TGCATTAATACCTAAC----- 735
QY 953 AsnThrSerLysGlyPheGluLeuHisPheSerSerPheGluLeuIleLysCysGluAsp 972
DB 736 -----AAATGACAGCCT 747
QY 973 ProGlyThrProLysPheGlyTyrLysValHisAspGlnGlyHisPhe-----AlaGly 990
DB 748 CCAATATGTGAAAT---GGAATATATGTAATCTGACACAGAGCTTATTTCTTAAAT 804
QY 991 SerSerValSerPheSerCysAspProGlyTyrSerLeuArgGlySerGluGluLeu 1010
DB 805 GAGTTGTGTGAGTTTGGTCTCAGCTTGGCTTTGTATGAAAGACCCCGCTGTGAAG 864
QY 1011 CysLeuSerGlyGluArgArgThrTyrAspArgProLeuProThrCysValAlaGlnCys 1030
DB 865 TGCAGAGCCCTGAACAA-----TGGAGACCGGAGACTACCAAGCTGCTCCAGGGATAT 918
QY 1031 GlyGlyThrValArgGlyValSerGlyGlnValLeuSerProGlyTyrProAlaPro 1050
  
```

```

DB 919 -----CAGCCACCTCCA 930
QY 1051 TyrGluHisAsnLeuAsnCysIleThrThrIleGluAlaGluAlaGlyCysThrIleGly 1070
DB 930 ----- 930
QY 1071 LeuHisPheLeuValPheAspThrGluGluValHisAspValLeuArgIleThrAspGly 1090
DB 931-----GATGTCCTGCAT----- 942
QY 1091 ProValGluSerGlyValLeuLeuLysGluLeuSerGlyProAlaLeuProLysAspLeu 1110
DB 942 ----- 942
QY 1111 HisSerThrPheAsnSerValValLeuGlnPheSerThrAspPheThrSerLysGln 1130
DB 942 ----- 942
QY 1131 GlyPheAlaIleGlnPheSerValSerThrAlaThrSerCysAsnAspProGlyIlePro 1150
DB 943 -----GCT 945
QY 1151 GlnAsnGlySerArgSerGlyAspSerTrpGluAlaGlyAspSerThrValPheGlnCys 1170
DB 946 GAGCGTACCCCAAGGACAGACAGACACTTTTCACTGGGACAGAGTCTTCACTACAGTCT 1005
QY 1171 AspProGlyTyrAlaLeuGlnGlySerAlaGluIleSerCysValIleIleGluAsnArg 1190
DB 1006 GAGCCCGGCTACGACATCAGAGGGGCTGCTGTATGCTGACACACCCAGGAGAC--- 1062
QY 1191 PhePheTrpGlnProSerProProThrCysIleAlaIleProCysGlyValAspLeuThrGly 1210
DB 1063 -----TGAGCCCTGACGCCCCACATGTAAGTGAATCTGTGATGACTTCAATGGAG 1116
QY 1211 Pro-----SerGlyValIleLeuSerProAsnTyrProGluProTyrProProGlyLys 1228
DB 1117 CAATCTCTTAATGCGCGTGTCTATTTCGA----- 1146
QY 1223 GluCysAspTrpLysValThrValSerProAspTyrValIleAlaLeuValPheAsnIle 1248
DB 1146 ----- 1146
QY 1249 PheAsnLeuGluProGlyTyrAspPheLeuHisIleTyrAspGlyArgAspSerLeuSer 1268
DB 1147 GTAATCTCCAGCTTGGAGCAAAAGTGAATTTGTTGTAT----- 1188
QY 1269 ProLeuIleGlySerPheTyrGlySerGlnLeuProGlyArgIleGluSerSerSerAsn 1288
DB 1189 -----GAAGATTTCATTAAGAAGC-----AGCTGCTAGT 1221
QY 1289 SerLeuPheLeuAla-----PheArgSerAspAlaSerValSerAsnAla 1303
DB 1222 TACTGTGCTTGGCTGGAGATGAAGACCTTGAATACAGATGTTCCAGTGTGTAACAA 1281
QY 1304 GlyPheValIleAspTyrThrGluAsnProArgGluSerCysPheAspProGlySerIle 1323
DB 1282 ATCTTT-----TCTCAAGCTCCCACTTAT 1308
QY 1324 LysAsnGlyThrArgValGlySerAspLeuLysLeu-----GlySerSerValThr 1340
DB 1309 CCTAATGGAGACACACAGAAACCTCTGAAAGCTTTCCCTTGGGAAACAGTAAT 1368
QY 1341 TyrTyrCysHis-----GlyLysTyrGluValAlaGluGlyThrSerThr 1354
DB 1369 TACACATGCGACCCCAACCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1428
QY 1355 LeuSerCysIleLeuGlyProAspGlyLysAspProValTyrAsnAsnProArgProValCys 1374
DB 1429 ATCCGCTGCACAAATGACCTCCAAAGGATGGGTTTGGAGACACCTCGCCCTCGCT 1488
QY 1375 ThrAlaProCysGlyGlyGlnTyrValGlySerAspGlyValValLeuSerProAsnTyr 1394
  
```

Db 1489 -----GGATTCTG----- 1497
 Qy 1395 ProGlnAsnThrSerGlyGlnLeuCybLeuThrValThrValProLeuAspTyr 1414
 Db 1497 ----- 1497
 Qy 1415 ValValPheGlyGlnPheAlaPhePheHisThrAlaLeuAsnValValGluValHis 1434
 Db 1497 ----- 1497
 Qy 1435 AspGlyHisSerGln-----HisSerArgLeuSerSerLeuSerGlySerHisThr 1452
 Db 1498 ----GGTCACTGTCAAGCCCGCATCATTTTCTGTTCCTCCAAAGTGAACCAACCAAT 1554
 Qy 1453 GlyLeuSerLeuProLeuAlaThrSerAsnGlnValLeuIleuYpPheSerAlaLeuGly 1472
 Db 1555 GCATGTGACTTTCCCATTTGGGACATCT-----TTAAAGTCGAAATGGCGCTT 1602
 Qy 1473 LeuAlaProAlaArgGlyPhe-----HisPheValTyrGlnAlaVal 1486
 Db 1603 GAGTACTACGGAGAGCCATTCTCTATGCATGTCTAGATAACCTGCTGTGTCAGT--- 1659
 Qy 1487 ProArg-----ThrSerAlaThrGlnCySerSerValProGluProArgTyrGlyLys 1504
 Db 1660 CCCAAAGATGCTGTAAACGTAAATCATGTAAACTCTCCAGATCCAGTGAATGGCATG 1719
 Qy 1505 ---ArgLeuGlySerAspPheSerValGlyAlaIleValArgPheGluCybAsnSerGly 1523
 Db 1720 GTGCATGTGATCAGACATCCAGGTTGGATCCAGAAATCAACTTCTTGACTACAGGG 1779
 Qy 1524 TyrAlaLeuGlnGlySerProGluIleGluCybLeuProValProGlyAlaLeuAlaGln 1543
 Db 1780 CACGAGCTCATGTGTCTACTCTATCTGCTGAATGATC---CTCTGGGCAATGCTGCCAT 1836
 Qy 1544 TrpAsnValSerAlaProThrCybVal---ValProCybGlyGlyAsnLeuThrGluArg 1562
 Db 1837 TGGAGCAGAAAGCCGCAATTTGTCAAGAAATCTTGTTGGGCTAACCCCCCATGGCC 1896
 Qy 1563 ArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeuAsnCybVal 1582
 Db 1897 AATGGAATTTTCATTAGCACACCAAGAAATTTTCATATGATCA----- 1944
 Qy 1583 TrpIleValValProGluGlyAlaGlyIleGlnIleGlnValValSerPheValThr 1602
 Db 1945 -----GTGGTACCTTACCGCTCC 1962
 Qy 1603 GluGlnAsnThrPheSerLeuGluValPheAspGlyAlaAspAsnThrValThrMetLeu 1622
 Db 1963 AATCCCGAAGCGGAGGAGAAAGTGTGAG-----CTTGTG 2001
 Qy 1623 GlySerPheSerGlyThrThrValProAlaLeuAsnSerThrSerAsnGlnLeuTyr 1642
 Db 2002 GGTCAGCCCTCC-----ATATACTGCACCAACAAATGAC----- 2034
 Qy 1643 LeuHisPheTyrSerAspIleSerValSerAlaIleGlyPheHisLeuGluTyrLeuThr 1662
 Db 2035 -----GATCAA 2040
 Qy 1663 ValGlyLeuSerSerCybProGluPro-----AlaValProSer----- 1675
 Db 2041 GTGGGATCTGGAGCGGCCCGGCTCAGTGCATTTACTTAACAATGACGCTCCA 2100
 Qy 1676 -----AsnGlyValIleThrGlyGluArg-----TyrLeuValAsnAspVal 1689
 Db 2101 AATGTGAAATATGAAATATGATATGACAAACAGAACTTATTTCTTAATGAAGATT 2160
 Qy 1690 ValSerPheGlnCybGluProGlyTyrAlaLeuGlnGlyHisAlaHisIleSerCybMet 1709
 Db 2161 GTGAGTTTAAAGTTCAGCTGCTTGTCTATGAAGAGACCCCGCTGAGACTGCCAG 2220
 Qy 1710 ProGlyThrValArgTyrPheAsnTyrProProProLeuCybIleAlaGlnCybGlyGly 1729
 Db 2221 ----GCCCTGAACAATGGAGCCGAGCTACCAAGCTGCTCCAGGGATATG----- 2268

Qy 1730 ThrValGluGluMetGluGlyValIleLeuSerProGlyPheProGlyAsnTyrProSer 1749
 Db 2268 ----- 2268
 Qy 1750 AsnMetAspCySerThrIleAlaLeuProValGlyPheGlyAlaHisIleGlnPhe 1769
 Db 2268 ----- 2268
 Qy 1770 LeuAsnPheSerThrGluProAsnHisAspTyrIleGluIleArgAsnGlyProTyrGlu 1789
 Db 2269 ----CAGCAGCTCCAGATGCTCTG----- 2289
 Qy 1790 ThrSerArgMetMetGlyArgPheSerGlySerGluLeuProSerSerLeuLeuSerThr 1809
 Db 2289 ----- 2289
 Qy 1810 SerHisGluThrThrValTyrPheHisSerAspHisSerGlnAsnArgProGlyPheLys 1829
 Db 2290 -----CATGCTGAGCGTACCCAAAG----- 2310
 Qy 1830 LeuGluTyrGlnAlaTyrGluLeuGlnGluCybProAspProGluProPheAlaAsnGly 1849
 Db 2311 ----GACAAAGCAACTTT----- 2325
 Qy 1850 IleValArgGlyAlaGlyTyrAsnValGlyGlnSerValThrPheGluCybLeuProGly 1869
 Db 2326 ----TCACCGGGAGAAATGTTTCTACACCTGTGAGCCGGC 2364
 Qy 1870 TyrGlnLeuThrGlyHisProValLeuThrCybGlnHisGlyThrAsnAlaAsnTyrAsp 1889
 Db 2365 TATGACTTCAGAGGCGGTGCTATGCGCTGCACA-----CCCAAGGAGACTGAGC 2418
 Qy 1890 HisProLeuProLysCybGluValProCybGlyValAsnIleThrSerSer-----Asn 1907
 Db 2419 CTTGCAGCCCCCAGATGTGAAGTGAATCTGTGATGACTTTCAGGCCCACTTCTTAAT 2478
 Qy 1908 GlyThrValTyrSerProGlyPheProSerProTyrSerSerSerGlnAspCybValTyr 1927
 Db 2479 GCGCGTGTCTATTTTCA----- 2496
 Qy 1928 LeuIleThrValProIleGlyHisGlyValArgLeuAsnLeuSerLeuGlnThrGlu 1947
 Db 2497 ----GTAAATCTCCAGCTTGAGCAAAAGT----- 2523
 Qy 1948 ProSerGlyAspPheIleThrIleTyrAspGlyProGlnGlnThrAlaProArgLeuGly 1967
 Db 2524 ----GATTTTGTGTGAT---GAAGGATTTCAATTAAAGC----- 2559
 Qy 1968 ValPheThrArgSerMetAlaLysLysThrValGlnSerSerSerAsnGlnValLeuLeu 1987
 Db 2560 ----AGCTCTGCTAGTTATTTGCTTG 2583
 Qy 1988 LysPheHisArgAspAlaAlaThrGlyIlePheAlaIleAlaPheSerAlaTyrPro 2007
 Db 2584 ----GCTGAAATGAAAGCTTTGGAATAGCAGTTCCA 2619
 Qy 2008 LeuThrLys-----CybProProProThrIleLeuProAsnAlaGluValThr 2024
 Db 2620 GTGTGTGAACAATCTTTGTCCAAAGTCCCAAGTATTTCTATGTGGAGACACACAGA 2679
 Qy 2025 GluAsn---GluGluPheAsnIleGlyAspIleValAlaGlyTyrArgCybLeuPro----- 2041
 Db 2680 AAACCTCTGGAAGCTTTCCTTTGAAAGAGATTAATTAACAATGCAAGCCCGACCCA 2739
 Qy 2042 ----GlyPheThrLeuValGlyAsnGluIleLeuThrCybLysLeu----- 2055
 Db 2740 GACAGAGGAGCAGCTTCACTCATTTGAGAGACACCACTCCGCTCACAAGTGAACCT 2739
 Qy 2056 ----GlyThrTyrLeuGlnPheGluGlyProProProIleCybGluVal----- 2070
 Db 2800 CAAGGAAATGGGGTTGG-----AGCAGCCCTGCGCTGTGAATTTCTGGT 2850

Db 4789 ACTAAT-----AAATGCACAGCTCCAGAGTT---GAAATGCAATTAGAGTACAGCA 4839
 Qy 2688 Ser-----PheapProgl1YThrValMetArgPheSerCyseGluAlaGlyHisVal 2704
 Db 4840 AACGAGATTCTTTTCCCTCACTGAGATCAATGATTAAATTAATGACCCCGGGTTTGTG 4899
 Qy 2705 LeuArgGlySerSerGluArgThrCyseGlnAlaAsnGlySerThrSerGlySerGlnPro 2724
 Db 4900 ATGTAGAGGTCCACACACTGTGCAGTGCAGACCAATGCAATGAGGGGCCCAAGCTGCA 4959
 Qy 2725 GluCyseGlyValIleSerCyseGlyAsnProgl1YThrProSerAsnAlaArgValAlaPhe 2744
 Db 4960 CACTGCTCCAGGGTGTGTCAGCCGCTCCAGAAATCTGCATGTGAGCATATCCCTA 5016
 Qy 2745 Ser-----AapGlyLeuValPheSerSerSerIleValIYrGluCyseArgGluGlyYr 2762
 Db 5017 AGCCATCAGACCAATTTCACCTGGGAGAGAGTTCTACAGCTGAGCCAGCTAT 5076
 Qy 2763 TyrAlaThrGlyLeuLeuSerArgHisCyseSerValaAsnGlyThrTrpThrGlySerAsp 2782
 Db 5077 GACCTCAGAGGGGCTGCGCTCTGTGCACCTGCACGCCCGGAGAGCTGAGCCCTGAGACC 5136
 Qy 2783 ProGluCyseLeuValIleAsnCyseGlyAsp-----Progl1YleProAlaAsnGlyLeu 2800
 Db 5137 CCTAGATGTACAGTGAATCCTGTGATGACTTCTGGGCAATCCCTCATGGCCGTGTG 5196
 Qy 2801 ArgLeuGlyAsnAspPheArgThrAsnLysThrValThrTrpGlnCyseValProgl1Yr 2820
 Db 5197 CTACTTCACCTTAATCTCCAGCTTGGGCAAGAGTCTTGTGTGGATGAAGGGTTC 5256
 Qy 2821 MetMetGluSerHisArgValSer-----ValLeuSerCyseThrLysAspArgThrTrp 2838
 Db 5257 CGATTAAAGGACAGCTGTGCTAGTCATGTGTCTTGGTGAAGAAA---GCCCTTGG 5313
 Qy 2839 AsnGlyThrLysProValCyseValaLeuMetCyseLysProProProLeuIleProAsn 2858
 Db 5314 AATAGCAGTGTCCAGTGTGGAACAAATCTTTGTCCAAATCTCCAGCTATCTTAT 5373
 Qy 2859 GlyLysValValGlySer-----AspPheMetTrpGlySerSerValThrTyrAla 2875
 Db 5374 GGGAGACACACAGAACTCCCTTGGAGATATTCCTATGAGAAAAGAAATCTTACGCA 5433
 Qy 2876 Cys-----LeuGluGlyYrGlnLeuSerLeu 2884
 Db 5434 TGCACACCCACCCAGACAGAGAGATGACCTTCAACTCATTGGG---GAGAGCTCCATC 5490
 Qy 2885 ProAlaValPheThrCyseGluGlyAsnGlySerTrpThrGlyLeuProGlnCys--- 2903
 Db 5491 CGTGCACAAAGTACCTCAAGGAGATGGGTTGGAGACAGCCCTGCGCTGTGA 5550
 Qy 2904 -----PheProValPheCyseGlyAspProgl1YValProSerArgIYArgArgGluAsp 2921
 Db 5551 CTTTCTGTCTCTGCTGCTGCCACATCCACCAGATCCAAAAGGGGCAATTAATGGA 5610
 Qy 2922 ArgGlyPheSerTrp-----ArgSerSerValSerPheSerCyseHisProProLeu 2938
 Db 5611 GGAACAGTATCTTATATCTTCTGGGATGACATACCTGACCTGTGAGCCCGGCTAC 5670
 Qy 2939 ValLeuValGlySerProArgArgPheCyseGlnSerAspGlyThrTrpSerGlyThrGln 2958
 Db 5671 CTGTAGTGGGAAAGGGCTCATTTTCTGTACAGACCAAGCAATCTGGAGCAATTGGAT 5730
 Qy 2959 ProSerCyseIleAspProThrLeuThrThrCyseAlaAspProgl1YValProGlnPheGly 2978
 Db 5731 CATTAATGCAAGAA-----GTAATGTAGCTTCCA---CTGTTATGAATGA 5778
 Qy 2979 IleGlnAsnAsn-----SerGlnGlyYrGlnValGlySerThrValLeuPheArg 2995
 Db 5779 ATCTCGAAGAGTTAGAAATGAAAAAATATATCATATGAGATATATGACTTTGAAG 5838
 Qy 2996 CyseGlnLysGlyYrLeuLeuGlnGlySerThrThrArgThrCyseLeuProAsnLeuThr 3015
 Db 5839 TGTGAAGATGGGTATCTCTGGAAGGCGAGTCCCTGAGCCAGTCCAGGGCGGATGACAGA 5898

Qy 3016 TrpSerGlyThrProPro 3021
 Db 5899 TGGAGC-----CTCTCT 5910
 RESULT 6
 5256642-1
 Patent No. 5256642
 APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; KONG,
 MINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
 H.; JAKSIDIS, SAVVAS; MARSH, HENRY C. JR.
 TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
 RECEPTOR 1 (CR1) AND A THROMBOCYTIC AGENT, AND THE METHODS OF
 USE THEREOF
 NUMBER OF SEQUENCES: 30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/588,128
 FILING DATE: 24-SEP-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 412,745
 FILING DATE: 26-SEP-1989
 APPLICATION NUMBER: 332,865
 FILING DATE: 03-APR-1989
 APPLICATION NUMBER: 176,532
 FILING DATE: 01-APR-1988
 SEQ ID NO.1:
 5256642-1
 LENGTH: 6951
 Alignment Scores:
 Pred. No.: 5,79e-93 Length: 6951
 Score: 1362.00 Matches: 537
 Percent Similarity: 32.83% Conservative: 266
 Best Local Similarity: 21.95% Mismatches: 833
 Query Match: 8.02% Indels: 810
 DB: 6 Gaps: 109
 US-10-016-248-2 (1-3104) x 5256642-1 (1-6951)
 Qy 798 CyseGluGluProGluValProAla-----TyrSerIleArgLysGlyLeuGlnPhe 814
 Db 337 TGTGTAATCCTCAGATCTCGTGAATGGCATGGTGCATGATCAAAAGGATCCAGTTC 396
 Qy 815 GlyValAlaGlyAspThrLeuThrPheSerCysePheProgl1YrArgLeuGluGlyThrAla 834
 Db 397 -----GATCCCAATTAATTAATCTTGTGACTAAAGATACCAATCATTTGCTCG 450
 Qy 835 ArgIleThrCys---LeuGlyGlyArgArgLeuThrPheSerProLeuProArgCys 853
 Db 451 TCTGCCACATGATCATCTCAGGTGATACGTCTATTGGGATATGAACCACTATTGT 510
 Qy 854 ---ValAlaGluCyseGlyAsnSerValThrGlyThrGlnGlyThrLeuLeuSerProAsn 872
 Db 511 GACAGAAATCTTGTGGGCTTACCCGCCCATGCACCAATGAGATTTTCATTAGCACCAAC 570
 Qy 873 PheProValAsnTrpAsnAsnAsnHisGluCysIleYrSerIleGlnThrGlnProGly 892
 Db 571 -----AGAGAAATTTTCACTATGATCACTGATGACCTTACCCCTGCATCTGGA 621
 Qy 893 LysGlyIleGlnLeuValaArgAlaPheGluLeuSerGluGlyAspValLeuLysVal 912
 Db 622 AGCGGA-----GGAGAAAGGTGTGAGCTT---GTGGTGAAGCC---TCCATA 666
 Qy 913 TyrAspGlyAsnAsnAsnSerAlaArgLeuLeuGlyValPheSerHisSerGluMet 932
 Db 667 TACTGACACCAATGATACGATCAA-----GTGGGATCTGAGAGCGGCCGCCCTCAG 720
 Qy 933 GlyValThrLeuAsnSerThrSerSerSerLeuThrLeuAspPheIleThrAspAlaGlu 952
 Db 721 TGCAATTATACCTAAC----- 735
 Qy 953 AsnThrSerLysGlyPheGluLeuHisPheSerSerPheGluLeuIleLysCyseGluAsp 972

Db ----- AAATGACGCCT 747
 Qy ProGlyThrProLysPheGlyTyrLysValHisAspGluGlyHisPhe----- AlaGly 990
 Db 748 CCAATATGCGAAAT---GGAATATGATCTGACACAGAACCTTATTTCTTAAAT 804
 Qy 991 SerSerValSerPheSerCysAspProGlyTyrSerLeuArgGlySerGluGluLeu 1010
 Db 805 GAAGTTGGAGTTAGTGTGACGCTGTCTTTGATGAGAACGCCCGGTGTAG 864
 Qy 1011 CysLeuSerGlyGluArgGlyThrTTPAspArgProLeuProThrCysValAlaGluCys 1030
 Db 865 TGGACGGCCTGAACAA---TGGAGCGGAGGCTTACCACTGCTCCAGGGTGTGT 918
 Qy 1031 GlyGlyThrValArgGlyGluValSerGlyGluValLeuSerProGlyTyrProAlaPro 1050
 Db 919 ----- CAGCCACTCCA 930
 Qy 1051 TyrGluHisAsnLeuAsnCysIleTTPThrIleGluAlaGluAlaGlyCysThrIleGly 1070
 Db 930 ----- 930
 Qy 1071 LeuHisPheLeuValPheAspThrGluGluValHisAspValLeuArgIleTTPAspGly 1090
 Db 931 ----- GATGCTCTGCAT--- 942
 Qy 1091 ProValGluSerGlyValLeuLeuGlyLeuSerGlyProAlaLeuProLysAspLeu 1110
 Db 942 ----- 942
 Qy 1111 HisSerThrPheAsnSerValValLeuGlnPheSerThrAspPheThrSerLysGln 1130
 Db 942 ----- 942
 Qy 1131 GlyPheAlaIleGlnPheSerValSerThrAlaThrSerCysAsnAspProGlyIlePro 1150
 Db 943 ----- GCT 945
 Qy 1151 GlnAsnGlySerArgSerGlyAspSerTrpGluAlaGlyAspSerThrValPheGlnCys 1170
 Db 946 GAGGTCACCAAGAGGACAGACACTTTTCACTGGCAGAGAAAGTCTTACAGTGT 1005
 Qy 1171 AspProGlyTyrAlaLeuGlnGlySerAlaGluIleSerCysValLysIleGluAsnArg 1190
 Db 1006 GAGCCCGCTACGACCTCAGAGGGGCTGCGTGTAGCGTCACACCCCGAGAGAC--- 1062
 Qy 1191 PheSerTrpGlnProSerProProThrCysIleAlaProCysGlyGlyAspLeuThrGly 1210
 Db 1063 -----TGGAGCCCTGCAGCCCCCACAATGTGAAGTGAATCTGTGATGACTTCATGGGC 1116
 Qy 1211 Pro-----SerGlyValIleLeuSerProAsnTyrProGluProTyrProProGlyLys 1228
 Db 1117 CAATTCTTAATGGCGGTGCTATTCCA----- 1146
 Qy 1229 GluCysAspTrpLysValThrValSerProAspTyrValIleAlaLeuValPheAsnIle 1248
 Db 1146 ----- 1146
 Qy 1249 PheAsnLeuGluProGlyTyrAspPheLeuHisIleTyrAspGlyArgAspSerLeuSer 1268
 Db 1147 GTAAATTTCCAGCTGGAGCAAAAGTGATTTGTGTGTAT----- 1188
 Qy 1269 ProLeuIleGlySerPheTyrGlySerGlnLeuProGlyArgIleGluSerSerSerAsn 1288
 Db 1189 -----GAAAGATTTCAATTAAAGGC-----AGCTTGCTAGT 1221
 Qy 1289 SerLeuPheLeuAla-----PheArgSerAspAlaSerValSerAsnAla 1303
 Db 1222 TACTGTCTTGTGGCTGAAATGAAAGCCTTTGGAATAGCAGTGTCCAGTGTGCAACA 1281
 Qy 1304 GlyPheValIleAspTyrThrGluAsnProArgGluSerCysPheAspProGlySerIle 1323
 Db 1282 ATCTTT-----TGTCAAAGTCTCCAGTTATT 1308

Qy 1324 LysAsnGlyThrArgValGlySerAspLeuLysLeu-----GlySerSerValThr 1340
 Db 1309 CTTAATGGAGACACACAGGAAACCTTGAAGTCTTTCCCTTTGGAAAAAGCATAAAT 1368
 Qy 1341 TyrTyrCysHis-----GlyGlyTyrGluValGluGlyThrSerThr 1354
 Db 1369 TACACATGGAGACCCCAACCAAGAGGAGACAGCTTGACCTCATTTGAGAGACACC 1428
 Qy 1355 LeuSerCysIleLeuGlyProAspGlyLysProValTTPAsnAsnProArgProValCys 1374
 Db 1429 ATCCGTGCACAAGTGAACCTCAAGGGAATGGGGTTGGAGCAGCCCTCCCTGCTGT 1488
 Qy 1375 ThrAlaProCysGlyGlyGlnTyrValGlySerAspGlyValValLeuSerProAsnTyr 1394
 Db 1489 -----GGAATTCG----- 1497
 Qy 1395 ProGlnAsnTyrThrSerGlyGlnIleCysLeuTyrPheValThrValProLysAspTyr 1414
 Db 1497 ----- 1497
 Qy 1415 ValValPheGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspValValGluValHis 1434
 Db 1497 ----- 1497
 Qy 1435 AspGlyHisSerGln-----HisSerArgLeuLeuSerSerLeuSerGlySerHisThr 1452
 Db 1498 ---GGTCACTGTCAAGCCCCCAGATCATTTTCTGTTGGCAAGTGTGAACCAACCAAT 1554
 Qy 1453 GlyGluSerLeuProLeuAlaThrSerAsnGlnValLeuIleLysPheSerAlaLysGly 1472
 Db 1555 GCATCTGACTTCCCATTTGGACATCT-----TTAAAGTACCAATGCCGTCT 1602
 Qy 1473 LeuAlaProAlaArgGlyPhe-----HisPheValTyrGlnAlaVal 1486
 Db 1603 GAGTACTAGGAGGACCATTTCTTATCATGCTAGATTAACCTGCTGCTCAAGT--- 1659
 Qy 1487 ProArg-----ThrSerAlaThrGlnCysSerSerValProGluProArgTyrGlyLys 1504
 Db 1660 CCCAAAGATGTCTGTAACGTAAATCATGTAAATCTCTCCGATGCTCAATGATGCATG 1719
 Qy 1505 ---ArgLeuGlySerAspPheSerValGlyAlaIleValArgPheGluCysAsnSerGly 1523
 Db 1720 GTGCATGTATGATCAGACATCCAGGTGGATCCAGAACTATTTCTGTACTAGAGG 1779
 Qy 1524 TyrAlaLeuGlnGlySerProGluIleGluCysLeuProValProGlyValAlaLeuAlaGln 1543
 Db 1780 CACCGACTCATTTGTCACTCATCTGCTGAATGTATC---CTTCGGGCAATGCTGCCAT 1836
 Qy 1544 TrpAsnValSerAlaProThrCysVal---ValProCysGlyGlyAsnLeuThrGluArg 1562
 Db 1837 TGGACACGAAACCCGCCCAATTTGTCAAGAAATCTTGTGGCTACCCCCACCATCGCC 1896
 Qy 1563 ArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeuAsnCysVal 1582
 Db 1897 AATGAGATTTCATTAGACCAACAGAGAAATTTCACTATGATCA----- 1944
 Qy 1583 TrpLysIleValValProGluGlyAlaGlyIleGlnIleGlnValValSerPheValThr 1602
 Db 1945 -----GTGGTACCTTACCGCTGC 1962
 Qy 1603 GluGlnAsnTrpAspSerLeuGluValPheAspGlyAlaAspAsnThrValThrMetLeu 1622
 Db 1963 AATCTGGAAGCGGAGGAGAAAGTGTAG-----CTTGCTG 2001
 Qy 1623 GlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerThrSerAsnGlnLeuTyr 1642
 Db 2002 GGTGAGCCCTCC-----ATATCTGCACACCAATGAC----- 2034
 Qy 1643 LeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHisLeuGluIleTyrLysThr 1662
 Db 2035 -----GATCAA 2040


```
Qy 1663 ValGlyLeuSerSerCysProGluPro-----AlaValProSer----- 1675
Db 2041 GTGGGCAATCTGGAGCGCCCGCCCTCAGTGCATTAATCAATAATGACCGCTCCA 2100
Qy 1676 -----AsnGlyValLeuSerGlyGluArg-----TyrLeuValAsnAspVal 1689
Db 2101 AATGTGAAATGGAATATGCTGATCTGACAAAGAGCTTAATTTCTTAAATGAAGTT 2160
Qy 1690 ValSerPheGlnCysGluProGlyTyrAlaLeuGlnGlyHisAlaHisIleSerCysMet 1709
Db 2161 GTGGAGTTAGTTCAGCTCGCTGCTTGTCAATGAAGACCCCGCTGTAAAGTCCAG 2220
Qy 1710 ProGlyThrValAlaGArgTyrAsnTyrProProProLeuCysIleAlaGlnCysGlyGly 1729
Db 2221 -----GCCCTGAACAAATGGAGCCGAGCTACCAAGCTCTCCAGGCTGTG----- 2268
Qy 1730 ThrValGluGluMetGluGlyValIleLeuSerProGlyPheProGlyAsnTyrProSer 1749
Db 2268 ----- 2268
Qy 1750 AsnMetAspCysSerTyrPylIleAlaLeuProValGlyPheGlyAlaHisIleGlnPhe 1769
Db 2268 ----- 2268
Qy 1770 LeuAsnPheSerThrGluProAsnHisAspTyrIleGluIleArgAsnGlyProTyrGlu 1789
Db 2269 -----CAACCACCTCCAGATCTCTG----- 2289
Qy 1790 ThrSerArgMetGlyArgPheSerGlySerGluLeuProSerSerLeuLeuSerThr 1809
Db 2289 ----- 2289
Qy 1810 SerHisGluThrThrValTyrPheHisSerAspHisSerGlnAsnArgProGlyPheLeu 1829
Db 2290 -----CATGCTGAGCTACCCAAAG----- 2310
Qy 1830 LeuGluTyrGlnAlaTyrGluLeuGlnGluCysProAspProGluProPheAlaAsnGly 1849
Db 2311 -----GACAAAGCAACTT----- 2325
Qy 1850 IleValArgGlyAlaGlyTyrAsnValGlyGlnSerValThrPheGluCysLeuProGly 1869
Db 2326 -----TCACCGCGGAGAGAGTGTCTACAGCTGTGAGCCCGGC 2364
Qy 1870 TyrGlnLeuThrGlyHisProValLeuThrCysGlnHisGlyThrAsnArgAsnTyrAsp 1889
Db 2365 TATGACCTCAGAGGGCGCTGCTATGCGCTGACA-----CCCCAGGAGACTGAGC 2418
Qy 1890 HisProLeuProLysCysGluValProCysGlyGlyAsnIleThrSerSer-----Asn 1907
Db 2419 CTGCGACGCCCCACATGTGAAGTGAATCCTGTGATGATCATGGGCCCAACTTCTAAT 2478
Qy 1908 GlyThrValTyrSerProGlyPheProSerProTyrSerSerSerGlnAspCysValTyr 1927
Db 2479 GGCCTGTGCTAATTTCCA----- 2496
Qy 1928 LeuIleThrValProIleGlyHisGlyValArgLeuAsnLeuSerLeuLeuGlnThrGlu 1947
Db 2497 -----GTAATCTCCAGCTTGAGCAAAAGT----- 2523
Qy 1948 ProSerGlyAspPheIleThrIleTyrAspGlyProGlnGlnThrAlaProArgLeuGly 1967
Db 2524 -----GATTTTGTGTTGAT--GAAGGATTTCAATTAAAGC----- 2559
Qy 1968 ValPheThrArgSerMetAlaLysLeuThrValGlnSerSerSerAsnGlnValLeuLeu 1987
Db 2560 -----AGCTCTGCTAGTATTGCTCTGTG 2583
Qy 1988 LysPheHisArgAspAlaAlaThrGlyGlyIlePheAlaIleAlaPheSerAlaTyrPro 2007
Db 2584 -----GCTGGAATGGAAGCCTTTGGAATAGCAGTTCCTCA 2619
Qy 2008 LeuThrLys-----CysProProProThrIleLeuProAsnAlaGluValValThr 2024
Db 2620 GTGTGTAACAATCTTTTGTGCCAAGTCTCCAGATTAATCTTAATGGAGACACACAGA 2679
Qy 2025 GluAsn---GluGluPheAsnIleGlyAspIleValArgTyrArgCysLeuPro----- 2041
Db 2680 AAACCTGTGAAAGTCTTCCCTTTGAAAGCAATTAATTAACATCGACGCCACCACCA 2739
Qy 2042 -----GlyPheThrLeuValGlyAsnGluIleLeuThrCysLysLeu----- 2055
Db 2740 GACAGAGGAGACGACTTCGACCTTATTTGAGAGAGACACCATCCGCTGCACAAGGACCT 2799
Qy 2056 -----GlyThrTyrLeuGlnPheGluGlyProProProIleCysGluVal----- 2070
Db 2800 CAAGGAATGGGGTTTG-----AGCAGCCCTGCCCTCGTGGAAATCTGGGT 2850
Qy 2071 HisCysProThrAsnGluLeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTyr 2090
Db 2851 CACTGTCAAGC----- 2862
Qy 2091 ProGlySerTyrProGlnPheGlnThrCysSerTyrPheValArgValGluProAspTyr 2110
Db 2863 -----CAGAT----- 2868
Qy 2111 AsnIleSerLeuThrValGluTyrPheLeuSerGluLysGlnTyrAspGluPheGluIle 2130
Db 2869 -----CATTTCTGTTCGCCAAGTTGAAGAACCAACCAACCAATGCA 2907
Qy 2131 PheAspGlyProSerGlyGlnSerProLeuLeuLysAlaLeuSerGlyAsnTyrSerAla 2150
Db 2908 TCTGACTTTCCTCAATGGAGACATCTTTAAAGTACCAAGCCGCTCGTACTACGGGAGG 2967
Qy 2151 ProLeuIleValThrSerSerSerAsnSerValTyrIleuArgTyrSerSer---AspHis 2169
Db 2968 CCATTTCTTACATGATGCTTAGATTAACCTGTC-----TGTCAAGTCCCAAGAT 3018
Qy 2170 AlaTyrAsnArgLysGlyPheLysIleArgTyrSerAlaProTyrCysSerLeuProArg 2189
Db 3019 GTCGTGAACGTAAATCA-----TGTAAATCCCTCCA 3051
Qy 2190 AlaProLeuHisGlyPheIleLeuGlnGlnThrSerThrGlnProGlyGlySerIleHis 2209
Db 3052 GATCCAGTGAATGAGCATGTGCATGTGATGATCAACATCCAGTGTGATCCAGATCAAC 3111
Qy 2210 PheGlyCysAsnAlaGlyTyrArgLeuValGlyHisSerMetAlaIleCysThrArgHis 2229
Db 3112 TATCTTGACTACAGGGACCGACCTCATTTGATCACTGCTGAATGATCTCTCA 3171
Qy 2230 ProGlnGlyTyrHisLeuTyrPserGluAlaIleProLeuCysGlnAlaLeuSerCysGly 2249
Db 3172 GGCATATCTGCCAT--TGGAGCAGAGAGCCGCATTTGTCAACGAATTCCTTGAGCG 3228
Qy 2250 LeuProGluAlaProLysAsnGlyMetValPheGly-----LysGluTyrThrVal 2266
Db 3229 CTACCCCAACCATCGCCAAATGGAATTCATTTAGCAACCAAGAGAGAAATTTCACTAT 3288
Qy 2267 GlyThrLysAlaValTyrSerCysSerGlnGlyTyrHis-----LeuGlnAla 2282
Db 3289 GGATCACTGTGTGACTTCCCGCTGCAATCTTGAAGACAGAGGAGAAAGTGTGAGCTT 3348
Qy 2283 GlyAlaGluAlaThrAlaGluCysLeu-----AspThrGlyLeuTyrPser--- 2297
Db 3349 GTGGGTAGCCCTCCATTAATCTGACACAGCATGACATCAAGTGGCATCTGAGCGGC 3408
Qy 2298 -----AsnArgAsnValProPro----- 2303
Db 3409 CCCGCCCTCAGTCAATTATTAACCAAAATGACAGCCTCCAATGTGAAATGAGATA 3468
Qy 2304 -----GlnCysVal 2306
Db 3469 TTGCTATCTGACACAGAGACTTATTTCTTAATGAAGTTGTGCACTTGAAGTGTGAG 3528
Qy 2307 Pro-----ValThrCys----- 2310
```

Db 3529 CCTGGCTTTTCATGAAAGACCCGCGTGTGAAGTGCAGGCCCTGAAACAAATGGAG 3588
 Qy 2211 ProbAbpValSerSerIleSer-----ValGluHisGlyArg 2322
 Db 3589 CCAAGTTACCAAGCTGCTCCAGGGGTGTGTACGCCGCTCCAGAAATCTCGCATGGAG 3648
 Qy 2223 TrpArgLeuIlePheGlnThrGlnThrGlnPheGlnIleAsnIleLeuMetLeuIleCysAsp 2342
 Db 3649 CATACCCCAAGCAGCATCAGACAACTTTTCACCTGGGAGAGAGTCTTACAGCTGTGAG 3708
 Qy 2243 ProGlyTyrTyrTyrThrGlyGlnArgValIleArgCysGlnIleAsnGlyLysTrpSer 2362
 Db 3709 CCTGGCATATGACCTCAAGGGGCTGCGTCTGCATGCACACCCAGGAGACTGTGAGC 3768
 Qy 2263 LeuGlyAspSerThrProThrCysArgIleIleSerCys-----GlyGluLeu 2378
 Db 3769 -----CCTGAAGCCCGAAGATGTGACAGTAAATCTGTGATGACTTCTTGGGCAATC 3822
 Qy 2379 ProIleProProAsnGlnIleArgIleGlyThrLeuSerVal---TyrGlyAlaThrAla 2397
 Db 3823 -----CCTCATGGCCGCTGTCTATTTCCACTTATCTCCAGCTTGGGGCAAGGTG 3873
 Qy 2398 IlePheSerCysAsnSerGlyTyrThrLeuValGlySerArgValArgGluCysMetAla 2417
 Db 3874 TCCTTTGTCTGTATGAAAGGGTTCGCTTAAAGGGCAGTTCCCTTACTGATGTGTCTTG 3933
 Qy 2418 AsnGly-----LeuTrpSerGlySerGlyValValArgCysLeuAlaGlyHisCysGly 2434
 Db 3934 GTTGGATGAAAGACCTTTGGAAATACAGTGTCTCTGTGTGTGAACATATCTTTTGCCA 3993
 Qy 2435 ThrProGluProIleValIleAsnGlyHisIleAsnGlyGlu-----AsnTyrSerTyr 2451
 Db 3994 AATCTCCAGCATCTCTTAATGGAGACACACAGAACTCCCTGTGAGATATCTCCAT 4053
 Qy 2452 ArgGlySerValValTyrGlnCysAsn-----AlaGlyPheArgLeu 2465
 Db 4054 GGAAGAAATATCTTACACATGTGACCCCAACCCACAGAGAGATGACCTTCAACCTC 4113
 Qy 2466 IleGlyMetSerValArgIleCysGlnGlnAspHis-----TrpSerGly 2481
 Db 4114 ATTGGGAGAGACCAATCCCGCTGCACAAAGTACCCTATGGGAATGGGGTTTGGAGACG 4173
 Qy 2482 LysThrProPheCys---ValProIleThrCysGlyHis----- 2493
 Db 4174 CTGCGCCCTCGCTGTGAATCTTCTGTGCTGTGTGCTGTAACCCAGAGCAATTT 4233
 Qy 2494 -----ProGlyAsnProValAsnGlyLeuThrGlnGlnAsnIlePheAsnLeu 2509
 Db 4234 CCAATTGCCAGCTCTACGATCCCAATTAATGACTTT-----GAGTTTCCAGTC 4281
 Qy 2510 AsnAspValValIlePheValCysAsnProGlyTyrMetAlaGluGlyAlaIleArgSer 2529
 Db 4282 GGGACATCTTTGAATTAATGATGCCCTCGGGATTTTGGAAATGTTCTCTATCTCC 4341
 Qy 2530 GlnCysLeuAlaSerGlyGlnTrpSerAspMetLeuProThrCysArgIleIleAsnCys 2549
 Db 4342 ---TGCTTAAGAAACTGTGTCTGTCAAGTGTGAAGACAACTGTGACGAATAATCATGT 4398
 Qy 2550 ThrAspProGlyIleGlnGlnIleAsnSerValArgGlnValIleAlaSerGlyProHisArg 2569
 Db 4399 GGAACCTCCACCAACCCCTCAATGAGATGTGCATTAATAACAGAT----- 4446
 Qy 2570 PheSerPheGlyThrThrValSerTyrArgCysAsnHisGlyPheTyrLeuLeuGlyTyr 2589
 Db 4447 ACAAGTTTGGATCAACAGTAAATTAATCTTGTAAAGAGGTTTGCATCATTTGGTTCC 4506
 Qy 2590 ProValLeuSerCysGlnGlyAspGly-----ThrTrpAspArgProArgProGln 2606
 Db 4507 CCAATCACTACTTGTCTGTCTGTCAAGCAATATGTACATGGATTAAGAGCACTATT 4566
 Qy 2607 CysLeuLeuValSerCysGlyHisProGlySerProProHisSerGlnMetSerGlyAsp 2626
 Db 4567 TGTGATATCATATCTTGT-----GAGCCACCTCCAAACATATCCATGGAGAC 4614

Qy 2627 SerTyrThr-----ValGlyAlaValValArgTyrSerCys--- 2638
 Db 4615 TTCTACAGAAACAATAGAAATCTTTTTCACAAATGAAACGGGTGTAACTTACAGTCCAC 4674
 Qy 2639 -----IleGlyLysArgThrLeuValGlyAsn 2647
 Db 4675 ACTGACCAAGATGAGAAACAGCTTTGAGCTTTGGAGGAGAAAGCGTCAATAT 4728
 Qy 2648 SerThrArgMetCysGlyLeuAspGlyHisTrpThrGlySerLeuProHisCysSerGly 2667
 Db 4729 TGCACAGCAAAAGATGATCAAGTGTGTGTGGAGACAGCCCTCCCTCGGTATATTCT 4788
 Qy 2668 ThrSerValGlyValCysGlyAspProGlyIleProAlaHisGlyIleArgLeuGlyAsp 2687
 Db 4789 ACTAAT-----AAATGACACAGCTCCAGAAATT---GAAATGCAATTGTAGATCAACAGA 4839
 Qy 2688 Ser-----PheAspProGlyThrValMetArgPheSerCysGlnIleGlyHisVal 2704
 Db 4840 AACAGAGTTTCTTTTCCCTCACTAGATCATCATTTAGATGTACGCCGGGTGTTC 4899
 Qy 2705 LeuArgGlySerSerGlyAspThrCysGlnIleAsnGlySerTrpSerGlySerGlnPro 2724
 Db 4900 ATGTGTGGGTCCCACTGTGCAGTGCACAGCAATGCGAGATGGGGGCCCAAGCTGCCA 4959
 Qy 2725 GluCysGlyValIleSerCysGlyAspProGlyThrProSerAsnAlaArgValAlaPhe 2744
 Db 4960 CACTGCTCCAGGAGT---TGTACGCCGCTCCAGAAATCTCTCATGTGATGACATACCTTA 5016
 Qy 2745 Ser-----AspGlyLeuValPheSerSerIleValTyrGlyCysArgGluGlyTyr 2762
 Db 5017 AGCCATCAGACAAACTTTTCACTGGGAGAGAGTTTACAGCTGTAGCCAGCTAT 5076
 Qy 2763 TyrAlaThrGlyLeuLeuSerArgHisCysSerValAsnGlyThrTrpThrGlySerAsp 2782
 Db 5077 GACCTCAAGGGGCTGCGTCTGTGCACAGCCCAAGGGAGACTGGAGCCCTGAAGCC 5136
 Qy 2783 ProGluCysLeuValIleAsnCysGlyAsp-----ProGlyIleProAlaAsnGlyLeu 2800
 Db 5137 CTTAGATGACAGTAAATCTGTATGATGATCTTCTGGGCAATCTCTCATGCGCGTGG 5196
 Qy 2801 ArgLeuGlyAsnAspPheArgTyrAsnLysThrValThrTyrGlnCysValProGlyTyr 2820
 Db 5197 CTACTTCCACTAATCTCAGCTTGGGCAAGAGTGTCTTTTGGATGAAGGTTTC 5256
 Qy 2821 MetMetGluSerHisArgValSer-----ValLeuSerCysThrLysAspArgThrTrp 2838
 Db 5257 CGATTAAAGGAGAGTCTGTAGTATGTGTCTTGGCTGGAGATGAAA---GCCCTTTGG 5313
 Qy 2839 AsnGlyThrLysProValCysLeuValIleLeuMetCysLysProProProLeuIleProAsn 2858
 Db 5314 AATAGCAGTTCCTCAGTGTGTGACAAATCTTTTCCAAATCTCCAGCTATCTCTTAAT 5373
 Qy 2859 GlyLysValValGlySer-----AspPheMetTrpGlySerSerValThrTyrAla 2875
 Db 5374 GGGAGACACACAGAACTCCCTTGGAGATATCTCTTAAGAAAGAAATATCTTAACGA 5433
 Qy 2876 Cys-----LeuGluGlyTyrGlnLeuSerLeu 2884
 Db 5434 TCGACACCCACCCACAGAGAGATGACCTTCAACCTCATTTGG---GAGAGCTTCATC 5490
 Qy 2885 ProAlaValPheThrCysGlnGlyAsnGlySerTrpThrGlyGluLeuProGlnCys--- 2903
 Db 5491 CCGTCACAGAGACCCCTCAAGAGATGGGTTTGGAGACAGCCCTCGCTGTGAA 5550
 Qy 2904 -----PheProValPheCysGlyAspProGlyValProSerArgGlyArgArgGluAsp 2921
 Db 5551 CTTTGTGTCTGTGCTGCTGCCACATCCACCAAGATCCAAAGGGGCATTATTCGA 5610
 Qy 2922 ArgGlyPheSerTyr-----ArgSerSerValSerPheSerCysHisProProLeu 2938
 Db 5611 GGAACGATATCTTATATCTTCTGGAGTGAACAATCACTACACTTGTGACCCCGGTAC 5670

Qy 2339 ValIeuValIGlySerProArgRgPheCyGlnSerAaspGlyThrTrpSerGlyThrGln 2958
 Db 5671 CTGTAAGGGGAAAGGCGCTTATTTCTGTACAGACCGGGAATCTGGAGCAATTGGAT 5730
 Qy 2959 ProSerCyAlIeaPProThrLeuThrThrCyAlaAaPProGlyValProGlnPheGly 2978
 Db 5731 CATATTGTCAAAGAA-----GTAAATTGTAGCTTCCCA---CTGTTATGAATGA 5778
 Qy 2979 IleGlnAaen-----SerGlnGlyThrGlnValIGlySerThrValIeuPheArg 2995
 Db 5779 ATCTCGAAGAGGATTAGAAATGAAATATATACATATGAGATTATGACTTGAAG 5838
 Qy 2996 CyGlnIuValGlyThrLeuLeuGlnGlySerThrThrArgThrCyLeuProAaenLeuThr 3015
 Db 5839 TGTGAAGATGGGTATCTCTGTGAAGGACGCTCCGAGCCAGCTCCAGCGGAGATGACAGA 5898
 Qy 3016 TrpSerGlyThrProPro 3021
 Db 5899 TGGGAC-----CCTCCT 5910

RESULT 7

5472939-1
 Patent No. 5472939
 APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
 MINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
 H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
 TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
 MEDIATED DISORDERS
 NUMBER OF SEQUENCES: 30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/138,825
 FILING DATE: 19-OCT-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 588,128
 FILING DATE: 24-SEP-1990
 APPLICATION NUMBER: 412,745
 FILING DATE: 26-SEP-1989
 APPLICATION NUMBER: 332,865
 FILING DATE: 03-APR-1989
 APPLICATION NUMBER: 176,532
 FILING DATE: 01-APR-1988
 SEQ ID NO:1:
 LENGTH: 6951
 5472939-1

Alignment Scores:
 Pred. No.: 5,79e-93 Length: 6951
 Score: 1362.00 Matches: 537
 Percent Similarity: 32.83% Conservative: 266
 Best Local Similarity: 21.95% Mismatches: 833
 Query Match: 8.02% Indels: 810
 DB: 6 Gaps: 109

US-10-016-248-2 (1-3104) x 5472939-1 (1-6951)

Qy 798 CyGlnIuValProGlnValProAla-----TySerIleArgGlyGlyLeuGlnPhe 814
 Db 337 TGTGTAATCTCCAGATCTGTGAATGGCATGGTGCATGATCAAGGATCAAGTTC 396
 Qy 815 G1yValIGlyAaPThrLeuThrPheSerCyAaPProGlyThrArgLeuGlnGlyThrAla 834
 Db 397 -----GGATCCCAAAATTAATATTTCTGTACTAAAGATACCAATCATTTGGTTCCTCG 450
 Qy 835 ArgIleThrCyAa-----LeuGlyGlyArgArgArgLeuTrpSerSerProIeuProArgCyA 853
 Db 451 TCTGCCCATGTCATCTCAGTGTGATATCTCATTTGGGATATGAAGAACCTATTGTGT 510
 Qy 854 ---ValIaGluCyGlnAaenSerValThrGlyThrGlnGlyThrLeuLeuSerProAaen 872
 Db 511 GACGAATTCCTGTGTGGGCTAACCCCAACCATCAACCAATGAGATTTCATTAGACCAAC 570
 Qy 873 PheProValAaenTyArgAaenAaenIleGluCyAlIeaPProGlnPheGlnPheGly 892

Db 571 -----AGAGAAATTTTCACTATGATGATCAGTGGTGAACCTACCGCTGCATCTCGA 621
 Qy 893 LyeGlyIleGlnIleuValAlaArgAlaPheGlnIleuSerGlnGlyAaPValIleuVal 912
 Db 622 AGCGGA-----GGAGAAAGTGTGTGAGCTT---GTGGGTGAGCCC---TCCATA 666
 Qy 913 TyArgPlyAaenAaenSerAlaArgLeuGlnValPheSerHisSerGluMetMet 932
 Db 667 TACTGACCAACGAAATGACGATCA-----GTGGGACTCTGGAGCGGCGCCCTCAG 720
 Qy 933 G1yValThrIeuAaenSerThrSerSerSerIeuTrpIeuAaPheIleThrAaPAlaGlu 952
 Db 721 TGCAATTATCTTAAC-----
 Qy 953 AaenThrSerGlyPheGlnIleuHisPheSerSerPheGlnIleuLysCyGlnAaP 972
 Db 736 -----AAATGACCGCT 747
 Qy 973 ProGlyThrProLysPheGlyTyArgValHisArgGlnGlyHisPhe-----AlaGly 990
 Db 748 CCNAATGTGAAAT---GGATATGTGTATCTGACACAGAGCTTATTTCTCTTAAT 804
 Qy 991 SerSerValSerPheSerCyAaPProGlyTySerIleuArgGlySerGlnIleuLeu 1010
 Db 805 GAAGTTGTGAGTTTAGGTGTGACCTGTCTTGTGTGAAGAGACCCCGCTGTAAG 864
 Qy 1011 CyLeuSerGlyGluArgGlnThrTrpAaPArgProIeuProThrCyValAlaGluCyA 1030
 Db 865 TGGCAGGCGCTGAACAAA-----TGGAGCGGAGCTACCAACGCTCTCCAGGGTATGT 918
 Qy 1031 G1yGlyThrValArgGlyGluValSerGlyGlnValIeuSerProGlyTyArgProAlaPro 1050
 Db 919 -----CAGCCACCTCA 930
 Qy 1051 TyArgHisAaenAaenCyAlIeTrpThrIleGlnAlaGluAlaGlyCyThrIleGly 1070
 Db 930 ----- 930
 Qy 1071 LeuHisPheLeuValPheAaPThrGluGluValHisAaPValIeuAaGlyIeTrpAaPgly 1090
 Db 931 -----GATGCTCTGCAT----- 942
 Qy 1091 ProValIuSerGlyValIleuLeuGluLeuSerGlyProAlaIeuProLysAaPleu 1110
 Db 942 ----- 942
 Qy 1111 HisSerThrPheAaenSerValIleuGlnPheSerThrAaPhePheThrSerIleGln 1130
 Db 942 ----- 942
 Qy 1131 GlyPheAlaIleGlnPheSerValSerThrAlaThrSerCyAaAaAaPProGlyIlePro 1150
 Db 943 -----GCT 945
 Qy 1151 GlnAaenGlySerArgSerGlyAaPSerTrpGluAlaGlyAaPSerThrValPheGlnCyA 1170
 Db 946 GAGCGTACCAAGGAGCAAGACAACTTTCACTGTGGAGAAAGTGTTCACGCTGT 1005
 Qy 1171 AaPProGlyTyArgAlaLeuGlnGlySerAlaGluIleSerCyValIleGluAaArg 1190
 Db 1006 GAGCCGGCTTACCACTCAAGGGCGTGCCTATAGCGCTGCACACCCAGGAGAC--- 1062
 Qy 1191 PhePheTrpGlnProSerProProThrCyAlIeaIleAProCyGlnGlyAaPleuThrGly 1210
 Db 1063 -----TGAGCCCTGAGCCCAACATCTGTAATCTGTGATGACTTCATGAGGC 1116
 Qy 1211 Pro-----SerGlyAlaIleIeuSerProAaenTyArgProIeuProTyArgProGlyLys 1228
 Db 1117 CAACTTCTTAATAGCGCTGTGCTATTTC----- 1146
 Qy 1229 GluCyAaPTrpLysValThrValSerProAaPTrpValIleAlaLeuValPheAaenIle 1248
 Db 1146 ----- 1146

OY		1249	PheAsnLeuGluProGlyTyrAspPheLeuHisIleTyrAspGlyValGAspSerLeuSer	1266
Dd	1147	GTAATCTCCACGGCTTGAGCAAAAGTAGATTGTGTGAT	-----	1188
OY	1269	ProLeuIleGlySerPheTyrGlySerClnLeuProGlyArgIleGlySerSerAsn	1388	
Dd	1189	-----GAAGCATTTCAATTAAAAGGC-----AGTCCTCGACT	1221	
OY	1289	SerLeuPheLeuAla-----PheArgSerAspAlaSerValSerAsnAla	1303	
Dd	1222	TACTGTGCTTGGCTGGAGAATGAAAAAGCTTTGGAAATGACAGTCTTCCAGTGTGAA	1281	
OY	1304	GlyPheValIleAspTyrThrGluAsnProArgLysSerCysPheAspProGlySerIle	1323	
Dd	1282	ATCTTT-----TCTCCAAGTCTCCAGTTAAT	1308	
OY	1324	LysAsnGlyThrArgValGlySerAspLeuLeu-----GlySerSerValThr	1340	
Dd	1309	CCTAATGGAGACACACAGAAAACCTCTGGAGACHTTCCCTTTGGAAAAAGATAAT	1368	
OY	1341	TyrrTyrrCysHis-----GlyGlyTyrGluValGlyGlyThrSerThr	1354	
Dd	1369	TACACATCGACCCCCACCACAGACAGAGAGAGAGCTTCGACTTCATTGGAGAGAAC	1428	
OY	1355	LeuSerCysIleLeuGlyProAspGlyLysProValTrpAsnAspProArgProValCys	1374	
Dd	1429	ATCGCGCTGCACAAGACACCTCTCAAGGAATGGGGTTTGAGACACACCGCGCCCTGGCT	1488	
OY	1375	ThrAlaProCysGlyGlyGlnTyrValGlySerAspGlyValIleLeuSerProAsnTyr	1394	
Dd	1489	-----GGAAATCTCG-----	1497	
OY	1395	ProGlnAsnTyrThrSerGlyGlnIleCysLeuTyrPheValThrValProLysAspTyr	1414	
Dd	1497	-----	1497	
OY	1415	ValValPheGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspValValGluValHis	1434	
Dd	1497	-----	1497	
OY	1435	AspGlyHisSerGln-----HisSerArgLeuLeuSerSerLeuSerGlySerHisThr	1452	
Dd	1498	---GGTCACTGTCAAAGCCCACGATCATTTTCTGTTGCCAAAGTGAACCAACCAAT	1554	
OY	1453	GlyGlnSerLeuProLeuAlaThrSerAsnGlnValLeuIleLysPheSerAlaLysGly	1472	
Dd	1555	GGATCTGACTTCCCATTTGGGAGACT-----TTAAAGTACGAATGCCGTCCT	1602	
OY	1473	LeuAlaProAlaArgGlyPhe-----HisPheValTyrGlnAlaVal	1486	
Dd	1603	GAGTACTACGGAGGCCAATTCCTATCACATGCTAGATAACCTGGCTGGTCAAGT--	1659	
OY	1487	ProArg-----ThrSerAlaThrClnCysSerSerValProGluProArgTyrGlyLys	1504	
Dd	1660	CCCAAAGATGCTGTAAACGTAAATCATGATAAAATCTCTCCAGATCCAGTAATGGCATG	1719	
OY	1505	--ArgLeuGlySerAspPheSerValGlyAlaIleValArgPheGluCysAsnSerGly	1523	
Dd	1720	GTCATCTGATCACAGACATCCAGGTGGATCAANAATCAATATTCTTGTAACAAGGG	1779	
OY	1524	TyrAlaLeuGlnGlySerProGluIleGluCysLeuProValProGlyAlaLeuAlaGln	1543	
Dd	1780	CACCGACTATTTGGTACTCATCTGCTGATATATC---CTTCGGGCAATGCTGCCAT	1836	
OY	1544	TrpAsnValSerAlaProThrCysVal---ValProCysGlyGlyAsnLeuThrGluArg	1562	
Dd	1837	TGGAGACAGAACCAGCGCAATTTGTCAAGAAATTCCTGTGGGCTACCCCCACCATGGCC	1896	
OY	1563	ArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeuAsnCysVal	1582	
Dd	1897	AATGGAGATTTCATTAGCACCAACAGAGAGATTTTCATCTATGATCA-----	1944	

QY	1563	TrpYslleValValP	ProGluGlyAlaGlyI	LeGlnIleGlnValVal	SerPheValThr	1602
Db	1945	-----	-----	-----	-----GTTGGACCTACCGCTCG	1962
QY	1603	GluGlnAntTrpAspSerLeuGluValPheAspGlyAla	AspAntThrValThrMetLeu	1622		
Db	1963	AATCGTAGAACCGAGGAGGAGAAAGGTTTGGAC	-----	-----CTTGTCG	2001	
QY	1623	GlySerPheSerGlyThrThrValProAlaLeuLeuAnSerThr	SerAnGlnLeuThr	1642		
Db	2002	GGTAGACCCCTCC	-----	-----ATATACGTGACCCACCAATGAC	2034	
QY	1643	LeuHisPheThrSerAPil	SerValSerAlaAlaGlyPheHisLeuGluThr	1662		
Db	2035	-----	-----	-----GATCAA	2040	
QY	1663	ValGlyLeuSerSerCysProGluPro	-----AlaValProSer	-----	-----	1675
Db	2041	GTGGGCATCTGAGAGGGGGCCGGCCCTCAGTGCATT	TTACTACCAAAATGACAGCCCTCCA	2100		
QY	1676	-----	-----	-----	-----	1689
Db	2101	AATGTGAAATGAAATATGTGTATCGACAACAGAGCTTAT	TTCTTAATAATGAGATT	2160		
QY	1690	ValSerPheGlnCysGluProGlyTyrAlaLeuGlnGlyHis	AlaHisIleSerCysMet	1709		
Db	2161	GTGGAGTTTAGTGTCAGCCGTGGCTTTGTCTATGAAAGACCCCGCCGCTGAGACTGGCAG	2220			
QY	1710	ProGlyThrValAlaArgTrpAntTy	ProProPheLeuCysIleAlaGlnCysGlyGly	1729		
Db	2221	-----GCCCTGAACAAATGGAGACCCGAGGCTACCAAGCTGCTCCAGGGATATG	-----	-----	-----	2268
QY	1730	ThrValGluGluMetGluGlyValI	IleLeuSerProGlyPheProGlyAsnThr	1749		
Db	2268	-----	-----	-----	-----	2268
QY	1750	AsnMetAspCysSerTrpValI	IleAlaLeuProValGlyPheGlyAlaHisIleGlnPhe	1769		
Db	2268	-----	-----	-----	-----	2268
QY	1770	LeuAsnPheSerThrGluProAsnHisAspTyrI	IleGluIleArgAnGlyProTyrGlu	1789		
Db	2269	-----	-----	-----	-----	2289
QY	1790	ThrSerArgMetMetGlyArgPheSerGlySerGluLeuProSerSerLeuLeuSerThr	1809			
Db	2289	-----	-----	-----	-----	2289
QY	1810	SerHisGluThrThrValTyrPheHisSerAspHisSerGlnAsnArgProGlyPhe	LeuAsp	1829		
Db	2290	-----	-----	-----	-----	2310
QY	1830	LeuGluTyrGlnAlaTyrGluLeuGlnGluCysProAspProGluProPheAlaAsnGly	1849			
Db	2311	-----	-----	-----	-----	2325
QY	1850	IleValArgGlyAlaGlyTyrAsnValGlyGlnSerValThrPheGluCysLeuProGly	1869			
Db	2326	-----	-----	-----	-----	2364
QY	1870	TyrGlnLeuThrGlyHisPheProValLeuThrCysGlnHisGlyThrAsnArgAsnTrpAsp	1889			
Db	2365	TATGACCTCAGAGGGGCTGCGTCTATGAGCGCTGCACA	-----	-----	-----	2418
QY	1890	HisPheLeuProLysCysGluValProCysGlyGlyValAsnIleThrSerSer	-----	-----	-----	1907
Db	2419	CCTGACAGCCCCACATGTAAGTAAATCTGTGATGACTTCATGAGGCCAACTTCTTAA	2478			
QY	1908	GlyThrValTyrSerProGlyPheProSerProTyrSerSerSerGlnAspCysValTrp	1927			
Db	2479	GGCCGCTGTCTATTTCCA	-----	-----	-----	2496
QY	1928	LeuIleThrValProIleGlyHisGlyValArgLeuAsnLeuSerLeuLeuGlnThrGlu	1947			

```

Db      2497  ---GTAATCTCCAGCTTGAGCAAAAGTG----- 2523
Qy      1948  ProSerGlyAspPheIleThrIleTyrAspGlyProGlnGlnThrAlaProArgLeuGly 1967
Db      2524  -----GATTTGTTTGAT---GAAGATTCAATTAAAGC----- 2559
Qy      1968  ValPheThrArgSerMetAlaLysLeuThrValGlnSerSerSerAsnGlnValLeu 1987
Db      2560  -----NACTCTGCTAGTATTGTCCTG 2583
Qy      1988  LysPheHisArgAspAlaAlaThrGlyGlyIlePheAlaIleAlaPheSerAlaTyrPro 2007
Db      2584  -----GCTGGAATGGAAGCCTTGGAAATGACAGTTCCA 2619
Qy      2008  LeuThrLys-----CysProProThrIleLeuProAsnAlaGluValAlaThr 2024
Db      2620  GTGTGTGAAACAATCTTTTGTCCAAAGTCTCCAGTATTCTTAATGGGAGACACAGCA 2679
Qy      2025  GluAsn---GluGluPheAsnIleGlyAspIleValArgTyrArgCysLeuPro----- 2041
Db      2680  AAACCTCTGGAAGTCTTCCCTTTGGAAAGACGAAATTAACACATGGACCCCAACCCA 2739
Qy      2042  -----GlyPheThrLeuValGlyAsnGlnIleLeuThrCysLeuLeu----- 2055
Db      2740  GACAGAGGACGAGCTTCGACCTCATTTGAGAGACAGACATCCGCTGCACAGACCT 2799
Qy      2056  -----GlyThrTyrLeuGlnPheGluGlyProProPheIleCysGluVal----- 2070
Db      2800  CAAGGAAATGGGGTTGG-----AGCAGCCTGCTCCCTGCTGGAATTCGGGT 2850
Qy      2071  HisCysProThrAsnGluLeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTyr 2090
Db      2851  CACTGTCAAGC----- 2862
Qy      2091  ProGlySerTyrProGlnPheGlnThrCysSerTrpLeuValArgValGluProAspTyr 2110
Db      2863  -----CCAGAT----- 2868
Qy      2111  AsnIleSerLeuThrValGluTyrPheLeuSerGluLysGlnTyrAspGluPheGluIle 2130
Db      2869  -----CATTTCTGTTGGCAGTGAACCAACCAACCAATGCA 2907
Qy      2131  PheAspGlyProSerGlyGlnSerProLeuLeuLysAlaLeuSerGlyAsnTyrSerAla 2150
Db      2908  TCTACCTTTCCCATTTGGGACATCTTAAAGTACAAATCCGCTCTGACTACGAGGAG 2967
Qy      2151  ProLeuIleValThrSerSerSerAsnSerValTyrLeuArgTyrPheSer---AspHis 2169
Db      2968  CCATTTCTATCACATGCTAGATTAACCTGCTC-----TGTCAAGTCCCAAGAT 3018
Qy      2170  AlaTyrAsnArgGlyGlyPheLysIleArgTyrSerAlaProTyrCysSerLeuProArg 2189
Db      3019  GTCTGTAAACGTAATCA-----TGTAAATCTCTCCA 3051
Qy      2190  AlaProLeuHisGlyPheIleLeuGlnThrSerThrGlnProGlyGlySerIleHis 2209
Db      3052  GATCCAGGAAATGGATGTCATGTGATCACAGACATCCAGTTGTGTCAGAAATCAC 3111
Qy      2210  PheGlyCysAsnAlaGlyTyrArgLeuValGlyHisSerMetAlaIleCysThrArgHis 2229
Db      3112  TATTTCTTACTACAGGACGACGACTCATTTGCTCATCTGCTGAATGATCTCTCCA 3171
Qy      2230  ProGlnGlyTyrHisLeuThrPheSerGluAlaIleProLeuGlyGlnAlaLeuSerCysGly 2249
Db      3172  GGCAATACCTGCCAT---TGAGACACGAAGCCGCCAATTTGTCAACGAATCTCTTGGG 3228
Qy      2250  LeuProGluAlaProLysAsnGlyMetValPheGly-----LysGluTyrThrVal 2266
Db      3229  CTACCCCAACCAATGCGCAATGAGATTTCATTAGACCAACAGAGAAATTTTCACAT 3288
Qy      2267  GlyThrLysAlaValAlaTyrSerCysSerGluGlyTyrHis-----LeuGlnAla 2282

```

```

Db      3289  GATCAGTGGTGAACCTACCGCTGCATCTTGGAAAGACAGAGGAGAAAGTGTTCGACT 3348
Qy      2283  GlyAlaGluAlaThrAlaGluCysLeu-----AspThrGlyLeuThrSer--- 2297
Db      3349  GTGGGTAGCCCTTCATATATCTGACACAGCAATGACATCAAGTGGGATCTGGAGCGC 3408
Qy      2298  -----AsnArgAsnValProPro----- 2303
Db      3409  CCGGCCCTCAGTGCATTAATACCTAACAAATGACAGCCTCCAAATGTGAAATGGAATA 3468
Qy      2304  ----- 2306
Db      3469  TTGATATGACAACAGAGCTTATTTCTTAATGAAGTTGGACTTAAAGTTCAG 3528
Qy      2307  Pro-----ValThrCys----- 2310
Db      3529  CCTGGCTTTGTATGAAGAAGACCCCGCGTGTGAAGTGCAGCGCCGAAACAAATGGAG 3588
Qy      2311  ProAspValSerSerIleSer-----ValGluHisGlyArg 2332
Db      3589  CCAAGTTACCAAGCTCTCCAGGCTGTGACCGCCCTCCAGAAATCTGCATGTGAG 3648
Qy      2323  TrpArgLeuIlePheGluThrGlnTyrGlnPheGlnAlaGlnLeuMetLeuIleCysAsp 2342
Db      3649  CATACCCCAAGCCATCAGACAACTTTCACTGCGCAGGAAGTGTCTACAGCTGTAG 3708
Qy      2343  ProGlyTyrTyrTyrThrGlyGlnArgValIleArgCysGlnAlaAsnGlyLysTrpSer 2362
Db      3709  CCTGGCTATACCTCAAGAGGGGCTGCTCTGCTGACATGCACACCCAGGAGATCTGAGC 3768
Qy      2363  LeuGlyAspSerThrProThrCysArgIleIleSerCys-----GlyGluLeu 2378
Db      3769  -----CCTAAGCCCGCAATGTGCAGTGAATCTCTGATGACTTCTGGGTCACTC 3822
Qy      2379  ProIleProProAsnGlnIleArgIleGlyThrLeuSerVal---TyrGlyAlaThrAla 2397
Db      3823  -----CCTAAGCCCGTGTCTATTTCCACTTATTTCCAGCTTGGGGCAAGGTG 3873
Qy      2398  IlePheSerCysAsnSerGlyTyrThrLeuValGlySerArgValAlaGluCysMetAla 2417
Db      3874  TCCCTTGTCTGTGTGAAGGGTTTCCTTAAAGGCGATTCCTGACTCATTTGCTTG 3933
Qy      2418  AsnGly-----LeuTrpSerGlySerGluValAlaArgCysLeuAlaGlyHisCysGly 2434
Db      3934  GTTGGAAATGAAACCTTTTGAATTAACAGAGTTCCTGTGTGAACATATCTTTTGCCA 3993
Qy      2435  ThrProGluProIleValIleAsnGlyHisIleAsnGlyGlu-----AsnTyrSerTyr 2451
Db      3994  AATCTCCAGCTATCTTAAATGGGAGACACAGAACTCCCTGTGAGATTCCTAT 4053
Qy      2452  ArgGlySerValValTyrGlnCysAsn-----AlaGlyPheArgLeu 2465
Db      4054  GGAAGAAGAAATATCTTACATGATGACCCCAACCCACACAGAGGATGACCTTCACCTC 4113
Qy      2466  IleGlyMetSerValArgIleCysGlnGlnAspHisHis-----TrpSerGly 2481
Db      4114  ATTGGGAGAGACCAATCCCGCTGCAAGTGAACCTCATGGGAATGGGGTTGGAGCAGC 4173
Qy      2482  LysThrProPheCys---ValProIleThrCysGlyHis----- 2493
Db      4174  CTTGCCCTCGCTGTGAATCTTCTGTCTGTGCTGTGCTGACGTGTAACCCAGACATTT 4233
Qy      2494  -----ProGlyAsnProValAsnGlyLeuThrGlnGlyAsnGlnPheAsnLeu 2509
Db      4234  CCATTTGCCAGCTCTACAGATCCCAATTAAGACTTT-----GAGTTTCCAGTC 4281
Qy      2510  AsnAspValValLysPheValCysAsnProGlyTyrMetAlaGluAlaIleArgSer 2529
Db      4282  GGGACATCTTTGAATTAATGAATGCGTCTGGGATTTTGGGAAATGTTCTATCTCC 4341
Qy      2530  GlnCysLeuAlaSerGlyGlnTyrPheSerAspMetLeuProThrCysArgIleIleAsnCys 2549
Db      4342  ---TGCTTAAGAAACTTGCTGTGCTGTCAGTGAAGTGAACAACTGTACAGAAATATCATGT 4398

```


Db 1548 ATCAGACACATCCAGTTGATCCAGAAATCACTATCTTGTACTACAGGACACGACCTC 1607
Qy 2219 ValGlyHisSerMetAlaIleCysThrArgHisProGlnIlyTyrrHisLeuTPSerGlu 2238
Db 1608 ATGGTCACACTCATCTGTAATGATATCTCTCAGGCAATCTGCCCAT--TGAGACAG 1664
Qy 2239 AlaIleProLeuGlyGlnAlaLeuSerCysGlyLeuProGlnAlaProIlyAsnGlyMet 2258
Db 1665 AAGCCGCAATTTGTCAACAAATCTCTGTGGGCTACCCCAACCATCGCAATGGAGAT 1724
Qy 2259 ValPheGly-----LysGluTyrrValGlyThrIlySalValIlyTyrrSerCysSer 2275
Db 1725 TTGATTAGACCAACAGAGAGAAATTTTCACTATGATGACAGTGTGACCTACCGCTGCAAT 1784
Qy 2276 GluGlyTyrrHis-----LeuGlnAlaGlyAlaGlnAlaThrAlaGluCysLeu 2291
Db 1785 CTGGAGACAGAGAGAGAAAGTGTGTGAGCTTGTGGGTGAGCCCTTCATATACCTGAC 1844
Qy 2292 -----AspThrGlyLeuTPSer-----Asp 2298
Db 1845 AGCAATGACATGACATGAGGACATCTGAGCGGCCGCCCTCAGTGCATTATACCTTAC 1904
Qy 2299 ArgAsnValProPro----- 2303
Db 1905 AATGACACGCTCCAAATGTGAAATATGAAATATGATGATCTGACAACAGAACTTATTT 1964
Qy 2304 -----GlnCysValPro----- 2307
Db 1965 TCCTTAATGAAAGTTGTGAGTTAGTTAGTGTACAGCTTGCTTGTCATGAAAGAACCCCG 2024
Qy 2308 ---ValThrCys-----ProAspValSerSerIleSer----- 2317
Db 2025 CGTGTGAAGTCCAGAGCCCTGAAACAAATGGAGACAGAGTTACCAAGCTGCTCCAGGGTG 2084
Qy 2318 -----ValGluHisGlyArgTyrArgLeuIlePheGlnIlyTyrr 2331
Db 2085 TGTACGCGGCTCCAGAAATCTGCAATGTCAGATGTCAGATACCCCAAGCATCAGACAACTTT 2144
Qy 2332 GlnPheGlnAlaGlnMetLeuIleCysAspProGlyTyrrTyrrTyrrGlnIlyAsn 2351
Db 2145 TCACCTGGGACAGAAAGTTCTTACAGCTGTGAGCTGTGCTATGACCTCAGAGGGGTGCG 2204
Qy 2352 ValIleArgCysGlnAlaAsnGlyLysTyrSerLeuGlyAspSerThrProThrCysArg 2371
Db 2205 TCTCTGACTGCAACCCCGAGAGACTGAGC-----CCTGAAGCCCGAAGATGTGCA 2258
Qy 2372 IleIleSerCys-----GlyLeuLeuProIleProProAsnGlyHisArgIle 2387
Db 2259 GTGAATCTGTATGATCTCTTGAGTCACTC-----CCTCATGGCCGTGTGCTA 2309
Qy 2388 GlyThrLeuSerVal---TyrGlyAlaThrAlaIlePheSerCysAsnSerGlyTyrr 2406
Db 2310 TTTCCACTATATCTCCAGCTTGGGGCAGAAAGTGTCTTGTCTGTATGAAAGGGTTTCCG 2369
Qy 2407 LeuValGlySerArgValArgGluCysMetAlaAsnGly-----LeuTPSerGly 2423
Db 2370 TTTAAAGGCGATTCCTGATGATCTGTGTGCTTGGTGTGAATGAGAACCTTTGGAATAC 2429
Qy 2424 SerGluValArgCysLeuAlaGlyHisCysGlyThrProGluProIleValAlaAsnGlyHis 2443
Db 2430 AGTGTTCCTGTGTGTAACATATCTTTGTCCAAATCTCCAGCTATCTTAATGGAGAG 2489
Qy 2444 IleAsnGlyGlu-----AsnTyrSerTyrArgIlySerValValIlyGlnCysAsn 2460
Db 2490 CACACAGAACTCCCTCTGAGATATCTCCATATGAAAGAAATATCTTACATATGAC 2549
Qy 2461 -----AlaGlyPheArgLeuIleGlyMetSerValAlaGlyIleCysGln 2474
Db 2550 CCCCACCCAGACAGAGGATGACCTTCAACCTCATTTGGGAGAGACCATCCGCTGACA 2609
Qy 2475 GlnAspHisHis-----TyrSerGlyLeuThrProPheCys---ValProIle 2489

Db 2610 AGTGAACCTCATGGGAATGAGGTTTGAGACAGCCCTGCGCTGTGAATCTTCTGT 2669
Qy 2490 ThrCysGlyHis-----ProGlyAsnProVal 2498
Db 2670 CCGTCTGTGACTGTGTAACCCCAAGACAGTTTCCATTGGCAATCTTACATCCCAAT 2729
Qy 2499 AsnGlyLeuThrGlnIlyAsnGlnPheAsnLeuAsnAspValIlyAspValCysAsn 2518
Db 2730 AATGACTTT-----GAGTTTCCAGTCCGGGACATCTTGAATATGAATGCCGT 2777
Qy 2519 ProGlyTyrrMetAlaGlnIlyAlaAlaArgSerGlnCysLeuAlaSerGlyGlnTPSer 2538
Db 2778 CCGGTGATATTTGGGAAATGTTCTTATCTCC---TGCCCTGAAATATGTTGCTGTGCTA 2834
Qy 2539 AspMetLeuProThrCysArgIlyIleIleAsnCysThrAspProGlyHisGlnIlyAsnSer 2558
Db 2835 AGTGTGAAGACAACTGTGAGAAATCATGTGACCTTCCACAGAACCTTCAATGGA 2894
Qy 2559 ValArgGlnValHisAlaSerGlyProHisArgPheSerPheGlyThrThrValSerTyr 2578
Db 2895 ATGTCGATATATAACACAGAT-----ACACAGTTTGGATCAACAGTATATAT 2942
Qy 2579 ArgCysAsnHisGlyPheTyrLeuLeuGlyThrProValLeuSerCysGlnIlyAspGly 2598
Db 2943 TCTGTATGAGAGGTTTGCATGATGTTCCCATCTACTACTTGTCTGTCTGCTCAGGC 3002
Qy 2599 -----ThrTyrAspArgProArgProGlnCysLeuLeuValSerCysGlyHisPro 2615
Db 3003 AATTAATGTCATGAGATAGAAAGACCACTATTTGTGAGATATATCTTGT----- 3053
Qy 2616 GlySerProProHisSerGlnMetSerGlyAspSerTyrThr----- 2629
Db 3054 ---GAGCCACTTCCAAACCATATCCAAATGAGAGACTTCTACAGCAACATAGAACCTTTT 3110
Qy 2630 ---ValGlyAlaValAlaArgTyrSerCys----- 2638
Db 3111 CCAATGGAACGCTGTGTAATCAACAGTGCACATGACAGACAGATGAGAAAGAGCTGTT 3170
Qy 2639 -----IleGlyLysArgThrIleValGlyAsnSerThrArgMetCysGlyLeuAspGly 2656
Db 3171 GAGCTTGTGGAGAACCGTCAATATAT-----TGACACAGAAAGATGATCAAGTTGGT 3224
Qy 2657 HisTPThrGlySerLeuProHisCysSerGlyThrSerValGlyValCysGlyAspPro 2676
Db 3225 GTTTGAGACAGCCCTCCCTCGGTGATATTTCTACTAAT-----AATGACAGCTGCA 3278
Qy 2677 GlyIleProAlaHisGlyIleArgLeuGlyAspSer-----PheAspProGlyThr 2693
Db 3279 GAAATT---GAAATGCAATTAGATACAGAGAAACAGAGAGTTTCTTCCCTCACTGAG 3335
Qy 2694 ValMetArgPheSerCysGlnAlaGlyHisValLeuArgIlySerSerGlnArgThrCys 2713
Db 3336 ATCATCAAGTTTAGATGTCAGCCGGGTTTGTCAATGAGAGGTCCACACTGTCAGTGC 3395
Qy 2714 GlnAlaAsnGlySerTyrSerGlySerGlnProGluCysGlyValIleIleSerCysGlyAsn 2733
Db 3396 CAGACCAATGAGAGATGGGGGCCCAAGCTGCCACACTGCTCCAGGTC---TGTACGCG 3452
Qy 2734 ProGlyThrProSerAsnAlaArgValAlaPheSer-----AspGlyLeuValPheSer 2751
Db 3453 CCTCCAGAAATCTCGATGATGAGCATACCTTAAGCATCAGAGCAACTTTTCACTGAG 3512
Qy 2752 SerSerIleValTyrGluCysArgGluGlyTyrTyrAlaThrGlyLeuLeuSerArgHis 2771
Db 3513 CAGGAAGTTTCTAAGCTGTGAGCCCAAGCTATGACCTCAGAGGGGCTCGCTTGCAC 3572
Qy 2772 CysSerValAsnGlyThrThrThrGlySerAspProGluCysLeuValIleAsnCysGly 2791
Db 3573 TGCAGCCCCCAGAGAGACTGAGAGCCCTGAAGCCCTTAATGTACACTGAATTCCTGTAT 3632
Qy 2792 Asp-----ProGlyIleProAlaAsnGlyLeuArgLeuGlyAsnAspPheArgTyrAsn 2809
Db 3633 GACTTCTGTGGCAACTCCCTCATGAGCGGTGTGTAATTCCTCACTTAATCTCCAGCTTGG 3692


```

Oy      2810  lyethrValThrrYrgLcYseValProglYTyrmctetGUseRhlAargValSer--- 2828
Dd      3693  GCAAAGGTCTCTTGGTTCGGCGAGAAGGGTTCCGATTAAAAAGCAGGTCTGTAAGTAT 3752
Oy      2829  ---ValleUserCysthlyeAspargThrTripannglYthrLyPproValCyblysAla 2847
Dd      3753  TGtGtCTTGCGCTGAATGAAA---GCCCTTGAAATACAGATGTTCAGTGTGTAACA 3809
Oy      2848  leumetCyblyPproProProleuIlleProanglYlyseValvalGlySer----- 2864
Dd      3810  ATCTTGTTCCAAAATCTTCAGCTATCCTTAATGGGAGACACAGAGAATCCCTTTGGA 3869
Oy      2865  AspphemetTrpgLYserSerValThrrTYrlaCYs----- 2876
Dd      3870  GATATTCCTTCATGAAAAAATAATCTTAACCATGCATCCACCACCCAGACAGAGAGAT 3929
Oy      2877  -----LeuglUglyTYrgLnleuSerleuProalaValPheThrCyeglUglyAsn 2893
Dd      3930  ACCTTCAACCTCATTTGGG---GAGAGCTCCATCCGCTCACAGAGTACCTCAAGGGAT 3986
Oy      2894  GlyserTPmThrglygluleuProglInys-----PheProValPhcyGlyasp 2910
Dd      3987  GGGGTGTGAGACAGCCCTGCCCTTCGCTGTAACTTTCTGTCTCGCTGCCACACAT 4046
Oy      2911  ProglYalProserArGlyArgarGlubAspargIlyPheSerTy-----Arg 2927
Dd      4047  CCACCCAAGATCCCAAAAGGGCATTAACATTGGAGACACAGATCTCTATATCTTCTGGG 4106
Oy      2928  SerSerValSerPheSerCyshieProProleuValleuValGlySerProArGArgPhe 2947
Dd      4107  ATGACAACTACACTACACTTGTGACCCCGGCTAACCTGTAGTGGGAAAGGGCTCATTTTC 4166
Oy      2948  CyeglnSerArpBglyThrtYrserylThrglnProserCySllaePProThrleuthr 2967
Dd      4167  TGTACAGACCAAGGGAATCTGGAGCCAATTGGAATCATTAATTCAAAGAA-----GTA 4217
Oy      2968  ThrCyalaAspProglYalProglInPheglYleglnAnaAsn-----SerGln 2984
Dd      4218  AATTGTAGCTTCCC--CTGTTATGAATGGAATCTCGAAGAGATTGAATGAAAAAA 4274
Oy      2985  GLYTYrgLnvalGlySerThrValleuPheArGcySglnlysgLYTYrleuLeunglnY 3004
Dd      4275  GTATATCATCTATGTAGATTATGTGACTTTGAAAGTGTGAAGATGGGTATATCTCGAAGGC 4334
Oy      3005  SerThrThrArGThrCySeLeuProAsnleuthrTpSerGlyThnProPro 3021
Dd      4335  AGTCCTCGAGACCCAGTGCAGAGCGGATGACAGATGGAC-----CTCTCT 4379

RESULT_9
5472939-3
Patent No. 5472939
APPLICANT: FEARON, DOUGLAS T.; KLIKSSTEIN, LLOYD B.; WONG, WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F., II; STEPHEN H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT MEDIATED DISORDERS
NUMBER OF SEQUENCES: 30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/138, 825
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 588,128
FILING DATE: 24-SEP-1990
APPLICATION NUMBER: 412,745
FILING DATE: 26-SEP-1989
APPLICATION NUMBER: 332,865
FILING DATE: 03-APR-1989
APPLICATION NUMBER: 176,532
FILING DATE: 01-APR-1988
SEQ ID NO.:3
LENGTH: 5420
5472939-3
```

Pred. No.:	1,1e-75	Length:	5420
Score:	1133.00	Matches:	402
Percent Similarity:	34.89%	Conservative:	197
Best Local Similarity:	23.41%	Mismatches:	620
Query Match:	6.67%	Indels:	498
DB:	6	Gaps:	77

US-10-016-248-2 (1-3104) x 547293-3 (1-5420)

QY	1494	CysSerSerValProGluProArgTrGlyLys--ArgLeuGlySerAspPheSerVal	1512
DB <td>156</td> <td>TGTAAGACTCTCCAGATCCAGTGAATGGCATGGTCATGTGATACAGACATCCAGTT</td> <td>215</td>	156	TGTAAGACTCTCCAGATCCAGTGAATGGCATGGTCATGTGATACAGACATCCAGTT	215
QY <td>1513</td> <td>GLYLALEVALAARGPHEGLUCYSANSESGLYTYRZLALEUNGILYSEPRGILULE</td> <td>1532</td>	1513	GLYLALEVALAARGPHEGLUCYSANSESGLYTYRZLALEUNGILYSEPRGILULE	1532
DB <td>216</td> <td>GGATCCAGAACTCACTATTCTTTGATCAACAGGACCGACCATTTGGTCACTACTCTGCT</td> <td>275</td>	216	GGATCCAGAACTCACTATTCTTTGATCAACAGGACCGACCATTTGGTCACTACTCTGCT	275
QY <td>1533</td> <td>GLUCYSLEUPROVALPROGLYVALALEUALAGINTRPANVALSERLAIPROTHRCSYVAL</td> <td>1552</td>	1533	GLUCYSLEUPROVALPROGLYVALALEUALAGINTRPANVALSERLAIPROTHRCSYVAL	1552
DB <td>276</td> <td>GAATGTATC---CTCTGGGGAATGCGCCATTTGAGACAGAAAGCCGCAATTTGTCA</td> <td>332</td>	276	GAATGTATC---CTCTGGGGAATGCGCCATTTGAGACAGAAAGCCGCAATTTGTCA	332
QY <td>1553</td> <td>--ValProCysGlyGlyAsnLeuThrGluArgGlyThrIleLeuSerProGlyPhe</td> <td>1571</td>	1553	--ValProCysGlyGlyAsnLeuThrGluArgGlyThrIleLeuSerProGlyPhe	1571
DB <td>333</td> <td>CGAATTCCTTTGGGGCTAACCCCCACCATCGCCATGAGATTTCATTAGCACCAAGCA</td> <td>392</td>	333	CGAATTCCTTTGGGGCTAACCCCCACCATCGCCATGAGATTTCATTAGCACCAAGCA	392
QY <td>1572</td> <td>PROGLUPROTYRLEUANSERLEUANSYCVALLTPRYSILEVALPROGLUGLYALA</td> <td>1591</td>	1572	PROGLUPROTYRLEUANSERLEUANSYCVALLTPRYSILEVALPROGLUGLYALA	1591
DB <td>393</td> <td>GAGAAATTTTCACTATGATCA-----</td> <td>413</td>	393	GAGAAATTTTCACTATGATCA-----	413
QY <td>1592</td> <td>GLYILEGILLEGINALVALSERPHEVALTHRGILUINSNTRPASPERSLEUGLVAL</td> <td>1611</td>	1592	GLYILEGILLEGINALVALSERPHEVALTHRGILUINSNTRPASPERSLEUGLVAL	1611
DB <td>414</td> <td>-----GTGGTACCTTACCGCTGCACATCCTGGAACGGAGGAGAAAGGTG</td> <td>458</td>	414	-----GTGGTACCTTACCGCTGCACATCCTGGAACGGAGGAGAAAGGTG	458
QY <td>1612</td> <td>PHEASPGLYLAASPAANTHRVALTHMETLEUGLYSERPHESEGLYTHRTHRVALPRO</td> <td>1631</td>	1612	PHEASPGLYLAASPAANTHRVALTHMETLEUGLYSERPHESEGLYTHRTHRVALPRO	1631
DB <td>459</td> <td>TTTGAG-----CTTGCGGTAGAGCCCTCC-----</td> <td>482</td>	459	TTTGAG-----CTTGCGGTAGAGCCCTCC-----	482
QY <td>1632</td> <td>ALAIEUENANSETRHNSERASNGILNEUTYLEUHSIPHETYRSEAPILSESERVAL</td> <td>1651</td>	1632	ALAIEUENANSETRHNSERASNGILNEUTYLEUHSIPHETYRSEAPILSESERVAL	1651
DB <td>483</td> <td>-----ATATCTGCACCGACCAATGAC-----</td> <td>503</td>	483	-----ATATCTGCACCGACCAATGAC-----	503
QY <td>1652</td> <td>SERLAALAGIPHEHSILEUGLUTYRLEUTHRVALLGLYLEUSERSEYSPROGLUPRO</td> <td>1671</td>	1652	SERLAALAGIPHEHSILEUGLUTYRLEUTHRVALLGLYLEUSERSEYSPROGLUPRO	1671
DB <td>504</td> <td>-----GATCAATGGGCATCTGGAGCGCGCGCCCT</td> <td>536</td>	504	-----GATCAATGGGCATCTGGAGCGCGCGCCCT	536
QY <td>1672</td> <td>-----AlaValProSer-----</td> <td>1681</td>	1672	-----AlaValProSer-----	1681
DB <td>537</td> <td>CAGTGCATTTAACTTAACAAATGACAGCCCTCCAATGTGAAATGAAATATGATGTCT</td> <td>596</td>	537	CAGTGCATTTAACTTAACAAATGACAGCCCTCCAATGTGAAATGAAATATGATGTCT	596
QY <td>1682</td> <td>GLUARG-----TYRLEUVALANSPALVALSERPHEGILCYSGIUPROGLIYTR</td> <td>1698</td>	1682	GLUARG-----TYRLEUVALANSPALVALSERPHEGILCYSGIUPROGLIYTR	1698
DB <td>597</td> <td>GACACAGAGACTTATTTCTTAAATGAAATGTGAGTTTACGTGACAGCCCTGGCTT</td> <td>656</td>	597	GACACAGAGACTTATTTCTTAAATGAAATGTGAGTTTACGTGACAGCCCTGGCTT	656
QY <td>1699</td> <td>ALAIEUNGILYHIALAHSLLESERCSYMERPROGLYTHRVALARGSTRIPASNTYR</td> <td>1718</td>	1699	ALAIEUNGILYHIALAHSLLESERCSYMERPROGLYTHRVALARGSTRIPASNTYR	1718
DB <td>657</td> <td>GTCATGAAGAGACCCCGCGGTGAAGTGCAG-----GCCCGAACAATGAGGACCG</td> <td>710</td>	657	GTCATGAAGAGACCCCGCGGTGAAGTGCAG-----GCCCGAACAATGAGGACCG	710
QY <td>1719</td> <td>PROBPROLEUCYSILLEALAGILCYSGIYGLYTHRVALGIUGIUMETGLUGLYVALILE</td> <td>1738</td>	1719	PROBPROLEUCYSILLEALAGILCYSGIYGLYTHRVALGIUGIUMETGLUGLYVALILE	1738
DB <td>711</td> <td>GAGCTACCAAGCTCTCCAGGGTATGT-----</td> <td>737</td>	711	GAGCTACCAAGCTCTCCAGGGTATGT-----	737
QY <td>1739</td> <td>LEUSERPROGLYPHEPROGLYDASNTYRPROSERASMETASERPYSERTRPLYLEALA</td> <td>1758</td>	1739	LEUSERPROGLYPHEPROGLYDASNTYRPROSERASMETASERPYSERTRPLYLEALA	1758
DB <td>737</td> <td>-----</td> <td>737</td>	737	-----	737
QY <td>1759</td> <td>LEUPROVALGIPHEGLYALAHSLILEGILNHEUANSNPHESETRHGIUPROASNHS</td> <td>1778</td>	1759	LEUPROVALGIPHEGLYALAHSLILEGILNHEUANSNPHESETRHGIUPROASNHS	1778
DB <td>738</td> <td>-----CAGCCACCTTCA</td> <td>749</td>	738	-----CAGCCACCTTCA	749
QY <td>1779</td> <td>ASPTYRILEGILILEARGASNGLYPROTYRGLUTHRSERARGMETMETGLYARGPHESE</td> <td>1798</td>	1779	ASPTYRILEGILILEARGASNGLYPROTYRGLUTHRSERARGMETMETGLYARGPHESE	1798

Db 750 GATGTCCTG----- 758
 Qy 1799 GlySerLeuLeuProSerSerLeuLeuSerThrsHisgluThrValTyrPheHis 1818
 Db 759 -----CAT 761
 Qy 1819 SerAspHisSerGlnAsnArgProGlyPheIleuGluTyrGlnAlaTyrGluLeuGln 1838
 Db 762 GCTGAGCGTACCCAAAG----- 779
 Qy 1839 GluCySerProAspProGluProPheAlaAsnGlyIleValArgGlyAlaGlyTyrAsnVal 1858
 Db 780 -----TCAACC 800
 Qy 1859 GlyGlnSerValThrPheGluCySerLeuProGlyTyrGlnLeuThrGlyHisProValLeu 1878
 Db 801 GGGAGAGAAAGTGTCTACAGCTGTGAGCCGGCTATACCTCAGAGGGGCTGCTATG 860
 Qy 1879 ThrCySerGlnHisGlyThrAsnArgAsnTyrAspHisProLeuProLysCySerGluValPro 1898
 Db 861 CGCTGCACA-----CCCCAGGAGACTGAGAGCCCTGCAGCCCATGTGAAGTGA 914
 Qy 1899 CySerGlyValAsnIleThrSerSer-----AsnGlyThrValTyrSerProGlyPhePro 1916
 Db 915 TCCTGTGATGACTTCATGGCGCAACTTCTTAATGGCCGTGCTGCTATTCCA----- 965
 Qy 1917 SerProTyrSerSerSerGlnAspCySerValTyrLeuIleThrValProIleGlyHisGly 1936
 Db 966 -----GTAAATCTCCAGCTTGAGCAAAA 989
 Qy 1937 ValArgLeuAsnLeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTyr 1956
 Db 990 GTG-----GATTTTGTGTGAT--- 1007
 Qy 1957 AspGlyProGlnGlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaLysLys 1976
 Db 1008 GAAGGATTTCAATTAAAGC----- 1028
 Qy 1977 ThrValGlnSerSerSerAsnGlnValLeuLeuLysPheHisArgAspAlaIleThrGly 1996
 Db 1029 -----AGCTCTGCTAGTATTGTGCTTG-----GCT 1055
 Qy 1997 GlyIlePheAlaIleAlaPheSerAlaTyrProLeuThrLys-----CysProPro 2013
 Db 1056 GGATGGAAGACCTTGGATAGAGAGTTCAGGTGTGACAAATCTTTTCCAAAG 1115
 Qy 2014 ProThrIleLeuProAsnAlaGluValValThrGluAsn--GluGluPheAsnIleGly 2032
 Db 1116 CCTCCAGATTATCTCTAATGGAGACACACAGAAAACCTCTGGAAGTCTTCCCTTGG 1175
 Qy 2033 AspIleValArgTyrArgCySerLeuPro-----GlyPheThrLeuVal 2046
 Db 1176 AAAGCACTAAATTACACATGACACCCCAACACAGAGGAGACGACTTCGACTCAT 1235
 Qy 2047 GlyAsnGluIleLeuThrCySerLysLeu-----GlyThrTyrLeuGlnPhe 2061
 Db 1236 GGAAGAGACACATCCGCTGCACACAGACCTCAAGGGAATGGGGTTGG----- 1286
 Qy 2062 GluGlyProProProIleCySerGluVal-----HisCySerProThrAsnGluLeuLeuThr 2079
 Db 1287 AGCAGCCCTGCCCTGCTGTGAAATTCTGGGTACACTGTCAAGCC----- 1331
 Qy 2080 AspSerThrGlyValIleLeuSerGlnSerTyrProGlySerTyrProGlnPheGlnThr 2099
 Db 1331 ----- 1331
 Qy 2100 CysSerTyrPheValArgValGluProAspTyrAsnIleSerLeuThrValGluTyrPhe 2119
 Db 1332 -----CCAGAT-----CATTTT 1343
 Qy 2120 LeuSerGluLysGlnTyrAspGluPheGlnIlePheAspGlyProSerGlyGlnSerPro 2139
 Db 1344 CTGTTCCTCAAGTTGAAACCAACCAATGATCTGACTTTCCTCATTTGGGACATCTTTA 1403

Qy 2140 LeuLeuLysAlaLeuSerGlyAsnTyrSerAlaProLeuIleValThrSerSerSerAsn 2159
 Db 1404 AAGTACGAATGCCCGTCCGATGACTACGGGAGCCATTTCTTATACATGCTTACATTAAC 1463
 Qy 2160 SerValTyrLeuArgTyrPheSerSer-----AspHisAlaTyrAsnArgLysGlyPheLysIle 2178
 Db 1464 CTGGTC-----TGCTCAAGTCCCAAGATGCTGTGAAGCGTAAATCA----- 1505
 Qy 2179 ArgTyrSerAlaProTyrCySerSerLeuProArgAlaProLeuHisGlyPheIleLeuGly 2198
 Db 1506 -----TGTAATACTCTCCAGATCCAGTAAAGGCAATGGCTGATCTGTG 1547
 Qy 2199 GlnThrSerThrGluProGlyLysSerIleHisPheGlyCysAsnAlaGlyTyrArgLeu 2218
 Db 1548 ATCACAAGATCCAGGTGGATCCAGATCACTATTCTTGATCTACAGGGACCGACATC 1607
 Qy 2219 ValGlyHisSerMetAlaIleCySerThrArgHisProGlnGlyTyrHisLeuThrPheSerGlu 2238
 Db 1608 ATTGGTCACTATCTGTGAATGATATCCCTCAGGGAATACGCCCAT---TGAAGCAG 1664
 Qy 2239 AlaIleProLeuCySerGlnAlaLeuSerCySerGlyLeuProGluAlaProLysAsnGlyMet 2258
 Db 1665 AAGCCGCCAATTGTTCACAGAAATTCCTGTGGGTACACCCCAACCATCGCCAAATGGAGAT 1724
 Qy 2259 ValPheGly-----LysGluTyrThrValGlyThrAlaValTyrSerCysSer 2275
 Db 1725 TTCATTAGACCAACACAGAGAAATTTTCACTATGATCTAGTGTGACCTACCGCTGCAT 1784
 Qy 2276 GluGlyTyrHis-----LeuGlnAlaGlyAlaGluAlaThrAlaGluCySerLeu 2291
 Db 1785 CTGGAGAGAGAGGAGAAAGGTGTTGAGCTTGAGGGGAGACCCCTCATATATCGACCC 1844
 Qy 2292 -----AspThrGlyLeuThrPheSer-----Asn 2298
 Db 1845 AGCAATGACATCAAGTGGCATCTGAGCGGCCCCCGCCCTCAGTGCATTAATACCTAAC 1904
 Qy 2299 ArgAsnValProPro----- 2303
 Db 1905 AATATGACGCTCCAAATGTGAAATGTGATCTGACACAGAAAGCTTATTT 1964
 Qy 2304 -----GlnCySerValPro----- 2307
 Db 1965 TCCTTAATAGAGTTGTGAGTTAGTGTGACGCTTGCTGTGATGAAGAACCCCGC 2024
 Qy 2308 ---ValThrCys-----ProAspValSerSerIleSer----- 2317
 Db 2025 CGTGTGAAGTGCAGGCGCTGAACAATGGAGCCAGAGATTACCAAGCTGCTCAGGGTG 2084
 Qy 2318 -----ValGluHisGlyArgTyrArgLeuIlePheGluThrGlnTyr 2331
 Db 2085 TGTCAAGCGCCTCCAGAAATCTGTGATGTGATGACATCCCAAGCATACAGACAACTTT 2144
 Qy 2332 GlnPheGlnAlaGlnLeuMetLeuIleCySerAspProGlyTyrTyrThrGlyGlnArg 2351
 Db 2145 TCACCTGGGACAGAGATGTTTACAGCTGTGAGCTGGGTATGACCTCAGAGGGCTGGC 2204
 Qy 2352 ValIleArgCySerGlnAlaAsnGlyLysTyrPheSerLeuGlyAspSerThrProThrCysArg 2371
 Db 2205 TCTTCGACTGCACACCCGAGAGAGATGAGAC-----CTGAAGCCCGGAGATGTGCA 2258
 Qy 2372 IleIleSerCys-----GlyGluLeuProIleProProAsnGlyHisArgIle 2387
 Db 2259 GTGAATCTGTGATGATCTTCTGGGTCACTC-----CTCATGCGCGGTGTGCTA 2309
 Qy 2388 GlyThrLeuSerVal---TyrGlyAlaThrAlaIlePheSerCysAsnSerGlyTyrThr 2406
 Db 2310 TTTCACCTTAATCTCAGCTTGAGGCAAAAGGTCTTGTCTGTGATGAAGGTTTGC 2369
 Qy 2407 LeuValGlySerArgValArgGluCysMetAlaAsnGly-----LeuThrSerGly 2423
 Db 2370 TTAAAGGGGAGTTCCTTATGATCATTTGTGTGTTGAATGAGAAAGCTTTGGAATTAAC 2429

Db 1000 GAAGAGTGAACCTTATCTCTTATTGGAGAGACACTCTCCGTTGTACCTGATATGTCAG 1059
Oy 2419 -----GlyLeuTrpSerGlySerGluValArgCys-----LeuAlaGlyHis 2432
Db 1060 AAGACTGGGACCTGGAGTGGCCCTGCCCAAGCTGTGAACCTTTCTACTTCGCGGTTGAG 1119
Oy 2433 CysGlyThrProGluProIleValAsnGlyHisIleAsnGly-----GluAsnTyrSer 2456
Db 1120 TGTCCACATCCCCCAACATCTTAAGAGCCGAATGCTATCTGGGACAGAAAGTCAATATACC 1179
Oy 2451 TyrArgGlySerValValTyrGlnCysAsnAlaGlyPheArgGluIleGlyMetSerVal 2470
Db 1180 TATATACGACACTGATATATTTGCTTGCATGTTTGGCTTCACCTTGAAAGGACGAACAA 1239
Oy 2471 ArgIleCysGlnGlnAspHisHisTrpSerGlyLysThrProPheCysValProIleThr 2490
Db 1240 ATCCGATCCATATGCCCAAGGACCATGGAGGCATGTGCACCATGTGTGAAG---GAA 1296
Oy 2491 CysGlyHisProGlyAsnProValAsnGlyLeuThrGlnGlyAsn-----GlnPhe 2507
Db 1297 TGCCAGGCGCCCTCCTAATCATCTCCTCAATGGGCAAAAGATAGACACATGATGCGCTTT 1356
Oy 2508 AsnLeuAsnAspValValLysPheValCysAsnProGlyTyrMetAlaGluGlyAla 2527
Db 1357 GACCCGTGAACATCTATATAATATATGCTGTAACTCTGCTATGCTGCTGGGAGAA 1416
Oy 2528 ArgSerGlnCysLeuAlaSerGlyGlnTrpSerAspMetLeuProThrCysArgIleIle 2547
Db 1417 TCCATACAGTACTCTCTGAGGGGGTGTGGACACCCCTGTACCCCAATGCACAAAGTGCA 1476
Oy 2548 AsnCys-----ThrAspProGlyHisGln----- 2555
Db 1477 GCGTGTGAAGTACAGAAAGGCACTTTGACAAACCAGACCAATTTGTTAGACA 1536
Oy 2556 -----GluAsnSerValArgGln----- 2561
Db 1537 GATGCAACTCTTCTGTGTGAGAGGTACAAATTAAAGTGGAGTGTATACAGAGTGT 1596
Oy 2561 ----- 2561
Db 1557 CAAGGCAAAATTCCTGGTTATGAGAGATTGCTTTGTAAAGAAATCACTGCCACCA 1656
Oy 2562 -----ValHisAlaSerGlyProHisArg-----PheSerPheGly 2573
Db 1657 CCCCCCTTATCTACAAATGGGGCAACACCGGAGTCTCTTAAGAAATTTCCATATGCA 1716
Oy 2574 ThrThrValSerTyrArgCysAsnHisGly-----PheTyrLeuLeu 2587
Db 1717 ACCACGGTCACTTACACATGTAAACCTGGCCAGAAAGAGATGCAATTCAGCCTCAT 1776
Oy 2588 GlyThrProValLeuSerCys-----GlnGlyAspGlyThrTTrpAspArgPro 2603
Db 1777 GAGAGAGACCACTTCGTTGTACAGCAATGATCAAGAAAGAGCAGCTGGAGTGGCCCT 1836
Oy 2604 ArgProGlnCys-----LeuLeuValSerCysGlyHisProGlySerProPro 2619
Db 1837 GCTCCCCATGTAAACTTCCCTCTCTGTCTGCAGTGTCTACATGTCCATATTCGAAT 1896
Oy 2620 HisSerGlnMetSerGlyAspSer-----TyrThrValGlyAlaValValArgTyrSer 2637
Db 1897 GGATACAGATATCTGGCAGAGAGACCCCAATTTCTCAATATGACATGTGACATTCAG 1956
Oy 2638 CysIleGlyLysArgThrLeuValGlyAsnSerThrArgMetCysGlyLeuAspGlyHis 2657
Db 1957 TGTATATAGTGAATTTACTTTGAAGGGCAGTGTCAATTCGTTGAAAGCTGATTAACA 2016
Oy 2658 TrpThrGlySerLeuProHisCysSerGlyThrSerValGlyValCysGlyAsp----- 2675
Db 2017 TGGATCTCGAAATCCAGTTTGTAAAAAGAAACA-----TGCACAGATGTGA 2067
Oy 2676 -----ProGlyIleProAlaHisGlyIleArgLeuGlyAspSerPheAspProGly 2692

Db	2068	CAGAGCTCTTCAAGAACTCCAGCT---	CGTTACAGTGG-----	2103
Qy	2693	ThrValMetArgPheSerCysGluAlaGlyHisValLeuArgGlySerSerGluArgThr	2712	
Db	2104	GAGCTAGTAAATTAAGCTCTGCCAAGATGGGTACAGATTGACCTGATTCATTCAGAT	2163	
Qy	2713	CysGln-----AlaGlnGlySerThrProSerGlyIserGlnProGluCysGlyValIleSer	2730	
Db	2164	TGTCAGATGCTGAAATATGAAATTTGGTCCAAAAGATTCACACTTTGTAAGATTATTCAC	2223	
Qy	2731	CysGluAspProGlyThrProSerAspAlaArg-----ValValPheSerAspGlyLeu	2748	
Db	2224	TGTCACCTCCACCAAGATGTTGTAATGGGAGACACACAGGATGATGGCGAAACTTT	2283	
Qy	2749	ValPheSerSerSerIleValIleArgIleCysArgGlnGlyIleArgIleValIleLeu	2768	
Db	2284	CTATATGGAAATGAAGTCTCTTATGAATGTGACCAAGATTCTATCTCTGGAGAGAA	2343	
Qy	2769	SerArgHisCys-----SerValGlnGlyThrThrProGlySerAspProGlu	2784	
Db	2344	AAATTGCAGTCACAGATGATTCTAAAGCATGATGATCTTGAGCCGGCCCTCCCAACAG	2403	
Qy	2785	CysLeu-----ValIleAsnCysGlyAspProGlyIleProAlaAsnGlyLeu	2800	
Db	2404	TGCTTAGCATCTCTCTCTGTACTCTGCTGCGCCCTTAATCCAGAGTC---AAACATGGGTAC	2460	
Qy	2801	ArgLeuGly-----AspArgPheArgGlyArgAspThrValIleThrGlyGlnCysVal	2817	
Db	2461	AAAGTCATTAATAACATCTTCGATATTCACACATGACATAGGTGATGTTGATCGCAT	2520	
Qy	2818	ProGlyIleMetMetGlnSerHisArgValSerValLeuSerCysThrIleAspArgThr	2837	
Db	2521	CGTGCTTCATCATGATATGCTAGTCGC-----GTGATTAGGTGTCATCTAGTAAACACA	2574	
Qy	2838	ThrAsnGlyThrIleProValCys-----LysAlaLeuMet---CysIleProProPro	2854	
Db	2575	TGGGTCCAGGTGGTGCACCTTGATATGAAACCTTCATAGGGTGTCCACCTCGCT	2634	
Qy	2855	LeuIleProAsnGlyLysValIleGlySerAsp-----PheMetTrpGlySerSer	2871	
Db	2635	AAGACCCCTTAACGGGAACCACTACGTGGTGAACAACATAGCTGATTTCTCCGGAATGTCA	2694	
Qy	2872	ValThrTyrAlaCysLeuGlnGlyIleGlnLeuSerLeuProAlaValPheThrCysGlu	2891	
Db	2695	ATCTGTACAGCTGTGACCAAGGCTACCTGCTGGGAGAGGACATCTCTTTGCACA	2754	
Qy	2892	GlyAsnGlySerThrPheGlyGlyLeuProGlnCysPheProValIlePheCysGlyAspPro	2911	
Db	2755	CATGAGGGAACCTGAGACCAACCTGCCCTCATTTGTAAGAAGGTAAACTGATGCTCA	2814	
Qy	2912	-----GlyValProSerArgGlyArgGlnAspArgGlyIlePheSerTyrArg	2927	
Db	2815	GCAGATATGATGGATATC---CAGAAAGGCTGGAAACCAAGAAATGATACAGTATGGA	2871	
Qy	2928	SerSerValIserPheSerCysHisIleProIleValIleValIleGlySerProArgArgPhe	2947	
Db	2872	GCTGTGTGAATCTCGAGTGTGAAGATGGAGATATGCTGGAAGGCAAGTCCCAAGCGAG	2931	
Qy	2948	CysGlnSerAspArgIleThrTrpSerGlyThrIleProSerCysIleAspProThrLeuThr	2967	
Db	2932	TGCCCATGTGATCAACAATGG-----AACCTCCCTCGCG	2967	
Qy	2968	ThrCysAlaAspPro-----	2972	
Db	2968	GTTTGCA---GATCCCGTCACTGCTCTGCTCTTGTGATGCTGACAGTTGATAC	3025	
Qy	2973	-----GlyValProGlnPheGlyIleGlnAAsnAsnSerGlnGly---	2985	
Db	3026	TTCCTACCTCTGATGTTGTCTATACCTTATACCGTATATCAAAACACAGAAACGCAATT	3085	
Qy	2986	-----TyrGlnValGlySerThrValLeuPheArgCysGlnIleValIleGlyLeuLeu	3002	
Db	3086	ATTATACAGATACACAGAAAGAGCTTTTACATTTTGAAGCGACAGAGTATATCTGT	3145	

```

QY      3003 GlnGlySerThrThrArgThrCysLeuProAsnLeuThrTrpSerGlyThrProProAsp 3022
DB      3146 TTGATCATACAAACCCAGCGCATGATCAGAAAGCAAACTGG----- 3187
QY      3023 CysValProHis 3026
DB      3188 TGTGTGCTCAT 3199

RESULT 11
US-09-285-385C-3
; Sequence 3, Application US/09285385C
; Patent No. 6579702
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S.
; APPLICANT: Scott, Ian C.
; APPLICANT: Thomas, Christina L.
; TITLE OF INVENTION: MAMMALIAN TOLLOID-LIKE GENE AND PROTEIN
; FILE REFERENCE: 960296.9611
; CURRENT APPLICATION NUMBER: US/09/285,385C
; CURRENT FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/111873
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/080550
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 4661
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (191)..(3229)
US-09-285-385C-3

Alignment Scores:
Pred. No.: 7.5e-49 Length: 4661
Score: 779.50 Matches: 344
Percent Similarity: 36.69% Conservative: 206
Best Local Similarity: 22.95% Mismatches: 508
Query Match: 4.59% Indels: 445
DB: Gaps: 55

US-10-016-248-2 (1-3104) x US-09-285-385C-3 (1-4661)
QY      1022 ProLeuProThrCysValAlaGluCysGlyGlyThrValArgGlyGluValSerGlyGln 1041
DB      239 CCGCTGCCCGCC-----GGTCTGAAGTACTGGGACCATTCGAAATGTC 283
QY      1042 ValLeuSerProGlyTyrProAlaProTyrGluHisAsnLeuAsnCysIleTPrThrIle 1061
DB      284 GCCTTGGACTACGGA-----GCCTTG 304
QY      1062 GluAlaGluAlaGlyCysThrIleGlyLeuHisPheLeuValPheAspThrGluGluVal 1081
DB      305 GAAGGGGAGGAGGCGACGAGCAGCAGCTGATTAC----- 340
QY      1082 HisAspValLeuArgIleTPrAspGlyProValGluSerGlyValLeuLeuGluLeu 1101
DB      341 CACGAC-----CCCTGCAAGCTGCTGCTCTTCTGGGGAGATATT 379
QY      1102 SerGlyProAlaLeuProLysAsp-LeuHisSerThrPheAsnSerValValLeuGlnPhe 1121
DB      380 -----GCTTGTGATGAAGATGATCTCAAGTTATTTCACATCCACAAGGCTTGAGAC 430
QY      1121 eSerThrAspPhePheThrSerLysGlnGlyPheAlaIleGln-PheSerValSerThr 1140
DB      431 TGGACCAAGCATTCATTGACAAACGAGCATGACATGAGAGGCTTGAGAGACATCT 490
QY      1141 AlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerGlyAspSerTrp 1160
DB      491 GCAAGGTGGCAAAAGHTACAGCC-----TTTAAAGCCAGCATCCAGGACCAAGAAAG 544

```

```

QY      1161 GluAlaGlyAspSerThrValPheGlnCysAspProGlyTyrAlaLeuGlnGlySerAla 1180
DB      545 GATGGCAAGATGCCCAACCATTTCTGCTTAACCTTGGAGACTCAAAACACACCGCTAA- 603
QY      1181 GluIleSerCysValIleGluAsn----- 1189
DB      604 GACCTTCTCTGCTCGAGTTGCAAGAGCTCAACCTCAAGACAGAGCGGATTTGGCCTGG 663
QY      1190 -----ArgPhePheThrGlnProSerProProThrCys 1200
DB      664 AAGGCTATTCCTTATGTCATTGAGAGAACTTACTGTGATCCAGAGGG-CCATTTTCA 722
QY      1201 IleAlaProCysGlyGly----- 1206
DB      723 AACAGGCCATGAGGACCTGGAGAAACACCTGTGTACCTTCGTAGAGAGACAGATG 782
QY      1207 -----AspLeuThrGlyProSerGlyValIleLeuSerProAsnTrp 1220
DB      783 AGGAAAGCTTCATTGATTCAGTTACAGGACCT-----GTGGTGTGTTCTTACTAGTGG 836
QY      1221 Pro-----GluProTyrProProGlyLysGluCysAsp----- 1231
DB      837 GAGCCCGAGAGGTGGCCCGCAGGCCATATTCATC-GGGAAAACTGTGACAAATTCCGC 895
QY      1232 -----TrpLysValThrValSerProAsp 1239
DB      896 ATTGGGCTCAGAGCTGGGCGCATGTGGTGGCTTGTGCATGAACACACTGGCCAGAC 955
QY      1240 TyrValIleAlaLeuValPheAsnIlePheAsnLeuGluProGly-----TyrAspPhe 1257
DB      956 CGAGACCAACATGATCATCATCATCAGAGAAACATCCAGCCAGTCAGAGATATATTC 1015
QY      1258 LeuHisIleTyrAspGlyArgAspSerLeuSerProLeuIleGlySerPheTyrGlySer 1277
DB      1016 TTTAAAAATGAAACCGGCGAGGTGAGC-----TCTTGGGAGAGACTTAC----- 1060
QY      1278 GlnLeuProGlyArgIleGluSerSerAsnSerLeuPheLeuAlaPheAspSerAsp 1297
DB      1061 -----GACTTCGACACATCATCATCATCAGTATCCCGGAACACC 1096
QY      1298 AlaSerValSerAsnAlaGlyPheValIleAspTyrThrGluAsnProArgGluSerCys 1317
DB      1097 TTCTCA-----AGAGAGTCTTCTTACAC-----ACCATCTCCCTCCGTGAGAC-- 1141
QY      1318 PheAspProGlySerIleLysAsnGlyThrArg-----ValGlySerAspLeuSerLeu 1335
DB      1142 -----GACATGGCGCTCAGGCCAACCATTTGGCCAAACGCTGGCGCTC 1183
QY      1336 GlySerSerValThrTyrCysHisGlyGlyTyrGluValGluGlyThrSerThrLeu 1355
DB      1184 AGTCAGGGAGATATAGCTCAAGCCAGCAAGCTGTCAAA----- 1222
QY      1356 SerCysIleLeuGlyProAspGlyLysProValTyrAsnAsnProArgProValCysThr 1375
DB      1223 -----TGCCCA 1228
QY      1376 AlaProCysGlyGlyGlnTyrValGlySerAspGlyValValLeuSerProAsnTrpPro 1395
DB      1229 GCA--TGGGGAGACACTACAGACACGACAGAAACTTTTCGGCACCTGTGTTCCCA 1285
QY      1396 GlnAsnTyrThrSerGlyGlnIleCysLeuTyrPheValThrVal-----ProLysAspTyr 1414
DB      1286 ATGAGCTACCCCTCTTATTCACACGCGCTGTGAGAGATCTCCGTACACCCAGGGGAAAG 1345
QY      1415 ValVal-----PheGlyGlnPheAlaPhePheHisThrAlaLeu-----AsnAspVal 1430
DB      1346 ATCATCTCAAAATTTACCTCATGATGATTTTAAAGACCCCTGTGCTGTGATCATAC 1405
QY      1431 ValGluValHisAspGlyHisSerGlnHisSerArgLeuLeuSerSerLeuSerGlySer 1450
DB      1406 GTGGAGATCCGGAGATGTACTGGAGAAAGGCCCTCTGTGGGAGGTTCTGTGCGCAT 1465

```

QY 1451 HisThrGlyGluSerLeuProLeuAlaThrSerAsnGlnValLeuIleLysPheSerAla 1470
 DB 1466 AAGATACCTGAGTCCCTT---GTCTCTCGGACAGCCGGCTCTGGGAAATTCGGTAGC 1522
 QY 1471 LysGlyLeuAlaProAlaArgGlyPheHisPheValTyrGlnAlaValProArgThrSer 1490
 DB 1523 AGCAGCAGCACCCTGGGCAAAAGGCTTCTTGCTGATATGAAGCCATG----- 1570
 QY 1491 AlaThrGlnCysSerSerValProGluProArgTyrGlyLysArgLeuGlySerAspPhe 1510
 DB 1570 ----- 1570
 QY 1511 SerValGlyAlaIleValAlaArgPheGluCysAsnSerGlyTyrAlaLeuGlnGlySerPro 1530
 DB 1570 ----- 1570
 QY 1531 GluIleGluCysLeuProValProGlyValaLeuAlaGlnITrPAsnValSerAlaProThr 1550
 DB 1570 ----- 1570
 QY 1551 CysValValProCysGlyGlyValAsnLeuThrGluArgArgGlyThrIleLeuSerProGly 1570
 DB 1571 -----TGTGGGGAGACATACCAAGATGCAAGGCGAGATTCAGTCTCCCAAC 1618
 QY 1571 PheProGluProTyrLeuAsnSerLeuAsnCysValITrPylsIleValValProGluGly 1590
 DB 1619 TACCTGACGACTACAGACTTCCCAAGGAATGTGTGTGAGAGATACAGTCCCGACGG 1678
 QY 1591 AlaGlyIleGlnIleGlnValValSerPheValThrGluGlnAsn-----Trp 1606
 DB 1679 TTCATGTGGGACTTACCTTCAGTCCCTTCAGATTCGAAGACATGACAGTGTGCAAT 1738
 QY 1607 AspSerLeuGlnValPheAspGlyValaAspAsnThrValIThrMetLeuGlySerPheSer 1626
 DB 1739 GACTATCTGGAAATCCGAGCGGTCCCAAGAGACAGCACCCTGATTTGGCCACTCTGT 1798
 QY 1627 GlyThrThrValProAlaLeuLeuAsnSerThrSerAsnGlnLeuTyrLeuHisPheTyr 1646
 DB 1799 GGCTACGAGAAAGCCGAGGCGGTAAATCCAGCGCTAACCGACTGTGGGTGAAGTTGTG 1858
 QY 1647 SerAspIleSerValSerAlaIleGlyPheHisPheGluTyrLysThrValGlyLeuSer 1666
 DB 1859 TCCACGCGCTCCATCAATAAAGCGGGCTTTCAGCCCAATTTC-----TTCAG 1906
 QY 1667 SerCysProGluProAlaValProSerAsnGlyValLysThrGlyGluArgTyrLeuVal 1686
 DB 1907 GAGGTGATGAGTCTCTCTGCGCAGACCATGT-----GGATGTGAGCAGCGCTGTGTA 1960
 QY 1687 AsnAspValValSerPhe-----GlnCysGluProGlyTyrAlaLeuGlnGlyHisAla 1704
 DB 1961 AACACACTCGGCACTACAGCTGTGCTGTGACCTGTGCTACGAATGGCTGCC----- 2014
 QY 1705 HisIleSerCysMetProGlyThrValArgITrPAsnTyrProProLeuCysIle 1724
 DB 2015 -----GACAAAGAAACATGTGAA 2032
 QY 1725 AlaGlnCysGlyGlyThrValGluGluMetGluGlyValIleLeuSerProGlyPhePro 1744
 DB 2033 GTGGCCGTGTGGTCTTCATTACCAAGTAAACGCAACCATCACAGCCCTGAGTGGCG 2092
 QY 1745 GlyAsnTyrProSerAsnMetAspCysSerITrPylsIleAlaLeuProValGlyPheGly 1764
 DB 2093 AAGAGAGTATCCCAACCAAGAACTGTGTCTGTGCGAGGTGGCTCCCGTGCATACGCC 2152
 QY 1765 AlaHisIleGlnPheLeuAsnPheSerThrGluProAsn-----HisAspTyr 1780
 DB 2153 ATCTCACTGCAAGTTCGAAGCTTTGAGCTGGAAGCAATGACGTGTAAATAGACTTC 2212
 QY 1781 IleGluIleArgAsnGlyProTyrGluThrSerArgMetMetGlyArgPheSerGlySer 1800
 DB 2213 GTAAAGGTGGCGAGTGTGTCTCCCAAGTTCACAGCTTCAAGCAATCTGTGGCTCC 2272
 QY 1801 GluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTyrPheHisSerAsp 1820

DB 2273 GAGACCCCGGAGGTATCATCATCGAGCAACAAACATGCGAGTGAATTCAGTCTGAC 2332
 QY 1821 HisSerGlnAsnArgProGlyPheLysLeuGlnTyrGlnAlaTyrGluLeuGlnGlyCys 1840
 DB 2333 AACACCGTCTCCAAAGCGAGGCTTCAGGCTCAGTCTTCTCA-----GACAAAGAACAGTGT 2389
 QY 1841 ProAspProGluProPheAlaAsnGlyIleValArgGlyAlaGlyTyrAsnValGlyGln 1860
 DB 2390 GCCAAAGAC-----AATGGCGGCTGCCAGCAGAGGTGTGTACACACCTTGGG 2437
 QY 1861 SerValThrPheGluCysLeuProGlyTyrGlnLeu-----ThrGlnHisProValLeu 1878
 DB 2438 AGTTACCTGTGCAGATGCAGAGACGGGTACCGATGCATGAGAACGGACAC----- 2488
 QY 1879 ThrCysGlnHisGlyThrAsnArgAsnITrPAsnHisProLeuProLysCysGluValPro 1898
 DB 2489 GACTGCAAA-----GAGGCTGGC 2506
 QY 1899 CysGlyGlyAsnIleThrSerSerAsnGlyThrValTyrSerProGlyPheProSerPro 1918
 DB 2507 TCGCCCTTACAAAGATCAGCAGTGCAGAGGGACCTGATGAGTCTTAACCTGGCCAGACAA 2566
 QY 1919 TyrSerSerSerGlnAspCysValITrPylsIleThrValProIleGlyHisGlyValArg 1938
 DB 2567 TACCCAGCCGGAAGAAATGTACTGGAACATTTCAACCGCAGCGCACAGGGTGA 2626
 QY 1939 LeuAsnLeuSerLeuLeuGlnIThrGluPro-----SerGlyAspPheIleThr 1954
 DB 2627 ATTACATTCAGTGAAGTTCGAGATTTGAGCAGACCAAGATGTGCTTATGACCATCTGGA 2686
 QY 1955 IleITrPAspGlyProGlnGlnIThrAlaProArgLeuGlyValIThrITrArgSerMetAla 1974
 DB 2687 CTGTACAGATGGAGACAGACAGCTTGGCCCCATCTTGCGCTTCTGCGGACAGAAAG 2746
 QY 1975 LysLysThrValGlnSerSerSerAsnGlnValLeuLeuLysPheHisAspAlaAla 1994
 DB 2747 CCGGATCCCGTGGTGGGAGCAGGACGACGACCTATCTCAGTCTTACCGGAGCGCTCA 2806
 QY 1995 ThrGlyGlyIlePheAlaIleAlaPheSerAlaTyrProLeuThrLysCysProPro 2014
 DB 2807 -----GTGACCGCGAAAGCCTTCAGGCTGTGACAGCAGACAGATG- 2848
 QY 2015 ThrIleLeuProAlaAlaGluValValIThrGluAsnGluGluPheAsnIleGlyAspIle 2034
 DB 2848 ----- 2848
 QY 2035 ValArgTyrArgCysLeuProGlyPheThrLeuValGlyAsnGluIleLeuThrCysLys 2054
 DB 2849 -----GGGGCAGGCTGAAGCT---GAGTACAGACCAACAAAGAG 2884
 QY 2055 LeuGlyIThrTyrLeuGlnPheGluGlyProProIleCysGluValHisCysProThr 2074
 DB 2885 CTCTATTCCCATGCCAGTTTGGG----- 2908
 QY 2075 AsnGluLeuLeuThrAspSerIThrGlyValIleLeuSerGlnSerTyrProGlySerTyr 2094
 DB 2909 -----GACAAACACTAC 2920
 QY 2095 ProGlnPheGlnIThrCysSerITrPylsValArgValGluProAspTyrAsnIleSerLeu 2114
 DB 2921 CCGAGCAGGCGCGCTGTGACTGGGTATGTGGCAGAAAGCGTTATGGCGTGAAGCTG 2960
 QY 2115 ThrValGluTyrPheLeuSerGlyLysGln-----TyrAspGluPheGluIle 2130
 DB 2981 ATATTCCGACCTTTGAAGTTGAGAGAGAGAGTACTGTGGCTACGACTTCATGAGAGCT 3040
 QY 2131 PheAspGlyProSerGlyGlnSerProLeuLeuValaLeuSerGlyAsnTyrSerAla 2150
 DB 3041 TATGATGGGTACGACAGACTGGACCCAGGCTGGCGCTTCTGTGTGC-----TCAGG 3094
 QY 2151 ProLeu---IleValIThrSerSerSerAsnSerValTyrLeuArgITrPAsnHis 2169

```

Db      3095 CCATTAGAGGAAATCTACTCCGCGAGACTCCGTAATGATCCGCTTCACACGAGAC 3154
Qy      2170 AlATyTAspAArgyVgslYPhelylLeArgTSeRlaProTyCySseRleuProAxy 2189
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3155 ACCATCAACAAGAAAGGCTTTCACGCGCGGTAC----- 3187
Qy      2190 AlAProLeuHieglYPhelIleAuglYglInThSeRInProglYglYserIleHs 2209
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3188 -----ACTAGACCAAGTTCCAAAGAGCGCTTGAC 3217
Qy      2210 Phe-glyCyAsnAlagly-----TyrArgLeuVa 2219
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3218 ATGAGGAAGTAGCGCCCTCAGCTTCGAAAGCAGAGAGACTGAGGTGTTTAAACCTT 3277
Qy      2219 lglYHiseRMeCAlaleCyRThr-----ArgHIsPro----- 2230
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3278 GCGAGTAGACAGCTCCCATGTACAGTGTGTTTTCACAAACAAACCAACTATGT 3337
Qy      2231 -----GInglYThIsleuTrpSerGlu----- 2238
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3338 TCTTGAACCTATATCTGGGTGACAGTACATGCTTGGCCAAAGGAGAGAGAGAG 3397
Qy      2239 -AlaIle-ProleuCySglnAlaleuSerCySglYleuProglYAlaProlyAsnGlyM 2258
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3398 GGCCTGTGTTCTGCTGTGCTGCTTATCAGTCGACGCTTCGAGAGAGTCAAGTTTGA 3457
Qy      2258 etValPheglYVgslYutYrThValglYThrlysaIaValYrSerCySserGluYr 2278
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3458 TGTTAGCGACCAACAGAGTACTTCTTCAATTTGCTCTTAAAGCTTGTCTCGTGGCT 3517
Qy      2278 YrHIsleuGlnAlaglyAlaglyAla-----ThralaglyCyS-----LeuAspT 2293
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3518 CCGAAGGCGACGCTGGGCGAAAGAGACAGCTTAGGGGAAAGACTGCTCTCTCACT 3577
Qy      2293 hrglyLeuTrpSerAenArgAsnValProProGlnCySValProvalThCyRProAspY 2313
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3578 CTGCGCCGCTTCATGATGAGTGCATGAGCCAGTGCCTGTGAGCTTGCGTGGCTCCCT 3633
Qy      2313 alSerSerIleSerValGlnHIsleglYArgrTrpArgleuIlePhegluThrglyGlnP 2333
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3634 -----CTAGCTACTGGGCA-TGAGAGATTATGATTGGCAGCGGCTTGCT 3678
Qy      2333 heglAlaglnLeuMeleuIleCySAspProglYrYrYrThrglyGlnArgValI 2353
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3679 GATACTGCAGGAGTGTGCTGACGCTTCGTGACGTTCTAC----- 3719
Qy      2353 leArgCySglnAlaAsnGlyLysTrpSerleuGlYAspSerThrProThCyArgIleI 2373
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3720 -----CAGAGTCAAGCCATCAAGAAACCGAGACCGTGCACAGAG----- 3761
Qy      2373 leSerCySglYgluLeuProIleProProAsnGlyHIsArgIleGly 2388
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3762 --AGTTAAGGTGGCATCCCA-----GGGCAAGAGAGGGG 3794

```

RESULT 12

US-08-991-408-1
Sequence 1, Application US/08991408
Patent No. 6008017

GENERAL INFORMATION:

APPLICANT: ARLETH, ANTHONY J.
APPLICANT: WILLETT, ROBERT N.
APPLICANT: ELSHOUBAGY, NABIL A.
APPLICANT: LI, XIOTONG

TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLOID-LIKE

TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: RATNER & PRESTIA

STREET: P.O. BOX 980

CITY: VALLEY FORGE

STATE: PA

COUNTRY: USA

ZIP: 19482

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,408
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/034,471
FILING DATE: 02-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: ATG-50038
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5145 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-991-408-1

Alignment Scores:
Pred. No.: 5,1e-48 Length: 5145
Score: 769.50 Matches: 245
Percent Similarity: 37.83% Conservative: 131
Best Local Similarity: 24.65% Mismatches: 331
Query Match: 4.53% Indels: 287
DB: 3 Gaps: 33

US-10-016-248-2 (1-3104) x US-08-991-408-1 (1-5145)

Qy      879 AsnAsnHIsGluCyelIeYrSerIleGlnThInProglYVgslYIleGlnLeuYs 898
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1023 GATACACCGTAACTATCAATAAGAGAAACATCCACCGCAGGTCA----- 1067
Qy      899 AlaArgAlaPhegluLeuSerGluGlyAspValLeuYValYrAspGlyAsnAsn 918
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1068 -----GAGTACATTTTCTGAGATGAGCTGAGAGCAAGCAAC 1106
Qy      919 SerAlaArgLeuLeuGlyVal---PheSerHIsSerGluMeCMeGlyValThrLeuAsn 937
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1107 TCA-----CTTGAGAGAAAGATATGATTTCGACAGTATCATGCACTATGCCAGGAAC 1157
Qy      938 SerThrSerSerSerleuTrpLeuAspPheIleThAspAlaglnAsnThrSerlyGly 957
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1158 ACCTCTCAAGGGGAGTGTTCGTGATACCATTTCTCCCTCCCT----- 1202
Qy      958 PhegluLeuHIsPheSerSerPhegluLeuIleYsCySgluAspProglY---ThrPro 976
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1203 -----GATGATATAGGCATACGTCT 1223
Qy      977 LysPheglYrYrlyValHIs---AspGluGlyHIsPheAlaglySerSerValSerPhe 995
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1224 GCATATGTGTAGGAGACCGCTCTAAGCAAGAGATATGCACAGGCA----- 1271
Qy      996 SerCySAspProglYrYrSerleuAspGlySerGluLeuLeuCySleuSerGlyglu 1015
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1271 ----- 1271
Qy      1016 ArgArgThrTrpAspArgProleuProThrCySValAlaglyCySglYglYrValArg 1035
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1272 -----AGAAAGCTGTATAGATGTCCGCA---TGTGAGAAACTCTACAA 1113
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      1036 glYgluValSerGlygluValleuSerProglYrYrProAlaProYrGluHIsAsnLeu 1055
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


Db 1314 ---GAATCAATGCAACCTTTCTCTCCAGATTTCCTCAATGAGCTACTTCTTACACA 1370
 Qy 1056 AsnCyseletprphrllleglualglualglucysrhrlllegluleuhsphleuval 1075
 Db 1371 CACGCAATCTGGAGAGTTTCTGTGACCCCAAGGAGAGATGTTTAAATTATTTACACG 1430
 Qy 1076 Pheasprhrngluval-----Hisapvalleuagilletpraspglpro 1091
 Db 1431 ATGATCTATACAGAGATGTTTGTCTGTATATATATGAAAGTAAGACGGGTAC 1490
 Qy 1092 ValgluserglyvalleuLeuLysgluleuSerglyProalaLeuProLysaspLeuHis 1111
 Db 1491 TGGAAATATCACTCTCTGTGAGATTCTGTGGGACAAATGCTGGAAGTTCTTACT 1550
 Qy 1112 SerthrPheanserValvalleuGlupheserThrAspPheThrSerLysglngly 1131
 Db 1551 TCTACACAGACAGAAATGTGATTGAGTTT----- 1580
 Qy 1132 PhealatlleglnPheSerValserThrAlathSerCysAsnAspProglylleProGln 1151
 Db 1580 ----- 1580
 Qy 1152 AsnGlySerArgSerglyAspSerglyProGlnAla---GlyaspSerThrValPheGlnCys 1170
 Db 1581 ---CGTACACAGACAGTAATGGGTAGGAAAGGCTTGCAGCTGCTATGAA--- 1628
 Qy 1171 AspProglyTyralaleuGlnGlySerAlaGluLeserCysValLyslleGluAsnArg 1190
 Db 1628 ----- 1628
 Qy 1191 PhehertprglnProSerProProThrCysillealProCysglYglYAspLeuThrGly 1210
 Db 1629 -----CGATCTGTGGAGGTGAGATCGTAA 1655
 Qy 1211 ProserGlyValilleuSerProAsnTyrrProGluProTyrrProProGlyLysglCys 1230
 Db 1656 AATGAGACAGATTCAGTCTCCATATATCTGATGACTATGCGCGAGTAAGAAATGT 1715
 Qy 1231 AspTrlyValThrValSerProAspTyrrValillealaleuValPheAsnillePheAsn 1250
 Db 1716 GTGGGAAATAAACAGTGTGAGAGCTACCACTCGGCGTACCTTTCAGTCTTTCAG 1775
 Qy 1251 LeuGluPro-----GlyTyrrAspPheLeuHisilleTyrrAspGlyArgAspSer 1266
 Db 1776 ATTTAAAGACATGACAAATTGCTTATGACTACCTGGAAGTTAGAGATGGAACCACTGAA 1835
 Qy 1267 LeuSerProLeuLleglySerPheTyrrGlySerglnLeuProGlyArgilleGluSerSer 1286
 Db 1836 AATAGCCCTTTGATAGGCGCTTCTGTGTTATGACAAACCTGAAAGCATTAAGTCTAAC 1895
 Qy 1287 SerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPheVal 1306
 Db 1896 TCCAAATCTTGTGATGAGATTGTTTCTGACGGAACCTGGAACAAAGCAGGTTTGTCT 1955
 Qy 1307 IleAspTyrrhrGluAsnProArgGluSerCysPheAspProGlySerilleLysAsnGly 1326
 Db 1956 GCTAACTTTTAAAGAG---GAAAGTAGTGTCCAACT----- 1994
 Qy 1327 ThrArgValGlySerAspLeuLysleuGlySerSerValThrTyrr-----Cys 1343
 Db 1995 GACCTGAGAGCTGTGACGACCATGTCTGAACACTGTGGCACTTCACTGCGCTGT 2054
 Qy 1344 HisGlyGlyTyrrGluValGluGlyThrSerThrLeuSerCysilleuGlyProAspGly 1363
 Db 2055 GAGCTGCGCTATGAG-----CTGGGCCACGAC--- 2081
 Qy 1364 LysProValTrpAsnAsnProArgProValCysThrAlaProCysGlyGlyGlnTyVal 1383
 Db 2082 -----AGAAGAGCTGTGAAGCTGCTTGTGTGGAGCTTCTTACC 2120
 Qy 1384 GlySerAspGlyValleuLeuSerProAsnTyrrProGlnAsnTyrrThrSerglyGlnlle 1403
 Db 2121 AAATTAACGACACATTAACACCCCTGTGCTGCGCCCAAGAGTACCTCTCTAATAAAGAC 2180
 Qy 1404 CysLeuTyrrPheValThrValProLysAspTyrrValValPheGlyGlnPheAlaPhePhe 1423
 Db 2181 TGTGTGGCAAGTGTGTGACCAACCCAGTACAGAAATTTCTGTGAAGTTTGATTTT 2240
 Qy 1424 HisThrAlaLeuAsnAspVal-----ValGluValHisAspGlyHisSer 1438
 Db 2241 GAATTTGAAGCAATAGAGTTTGCAAAATATGATTAATGTGAGATCTGAGAGTGTCTTCC 2300
 Qy 1439 GlnHisSerArgLeuLeuSerSerLeuSerglySerHisThrGlyGluSerLeuProLeu 1458
 Db 2301 TCTGAGCTTAACCTGATGCAAAATTTCTGTGGCGTCAAGTGCCTGAAGATC---ACA 2357
 Qy 1459 AlaThrSerAsnGlnValleuLysPheSerAlaLysGlyLeuAlaProAlaArgGly 1478
 Db 2358 TCCAGATTCAACAATATGAAATTAATTAATTCAGACATATCTGATTCAGAAAGGCG 2417
 Qy 1479 Phe-----HisPheValTyrrGlnAlaValProArgThrSerAlaThrGlnCysSerSer 1496
 Db 2418 TTCAAGACATTTTTC-----TCAGACAAAGATGATGCTCT--- 2456
 Qy 1497 ValProGluProArgTyrrGlyLysArgLeuGlySerAspPheSerValGlyAlaileVal 1516
 Db 2457 ---AAGATATGATGATGATGTCAGACGAATGTGTCAACAGATGGAGCTATAG 2510
 Qy 1517 ArgPheGlnCysAsnSerglyTyrrAlaLeuGlnGlySer-----ProGluIle 1532
 Db 2511 ---GTCAATGCGGTAAATGATTGTGCTACATGACATTAACATGATTCAGAAAGAGCT 2567
 Qy 1533 GluCysLeuProValProGlyAlaLeuAlaGlnTrpAsnValSerAlaProThrCysVal 1552
 Db 2568 GAGTGT-----GAACAGAAAGTCCACAGCTCAACT----- 2597
 Qy 1553 ValProCysGlyGlyAsnLeuThrGluArgArgGlyThrilleuSerProGlyPhePro 1572
 Db 2598 -----GGCTCATCATCAGTCCCACTGCGCCA 2624
 Qy 1573 GluProTyrrLeuAsnSerLeuAsnCysValTrpLysilleValValProGluGlyAlaGly 1592
 Db 2625 GACAAAGTACCAAGCAGAAAGAAATGACACTTGGGAAATGACGCCACTCTGTGCCACGA 2684
 Qy 1593 IleGlnilleGlnValValSerPheValThrGlnGlnAsn-----TrpAspSer 1608
 Db 2685 ATCAAAATTAACCTTTAGTGAATTTGAGATTGAGACAGCATTAAGATGCTTATGACAC 2744
 Qy 1609 LeuGluValPheAspGlyAlaAspAsnThrValThrMetLeuGlySerPheSerglyThr 1628
 Db 2745 TTGAAGTATTTGATGAGAAACAGAAAGTACCCGATTTCTGACGACTATGTGGCAAC 2804
 Qy 1629 ThrValProAlaLeuLeuAsnSerThrSerAsnGlnLeuTyrrLeuHisPheTyrrSerAsp 1648
 Db 2805 AAGATACCAATCCCTTGTGTGCTACTGGAATAAATGTTTGTGCGTTTGTTCGAT 2664
 Qy 1649 IleSerValSerAlaLaglyPheHisleuGluTyrrLysThrValGlyLeuSerSerCys 1668
 Db 2865 GCATCTGTTTAAAGAAAGGCTTCAAGCCACACATTTCA----- 2906
 Qy 1669 ProGluProAlaValProSerAsnGlyValLysThrGlyGluArgTyrrLeuValAsnAsp 1688
 Db 2906 ----- 2906
 Qy 1689 ValValSerPheGlnCysGluProGlyTyrrAlaLeuGlnGlyHisAlaHisIleSerCys 1708
 Db 2906 ----- 2906
 Qy 1709 MetProGlyThrValArgArgTrpAsnTyrrProProLeuCysillealGlnCysGly 1728
 Db 2907 -----GAGTGTGCG 2915
 Qy 1729 GlyThrVal-----GluGluMetGluGlyValilleuSerProGlyPheProGly--- 1745
 Db 2916 GAGCAATTAAGACAGAAATCAAAACCAAGATCTGTACTCAATGTCTGATTTGGTAT 2975

QY 1746 ---AsnIYrProSerAenMeuAaspCySeSerTrpIysIleAlaLeuProValIGlyPheGly 1764
 DB 2976 AACAACTAACCCAGACAGCTTGACTGTGAATGGCTATTAGTATCAGAACGGGGCTCCGA 3035
 QY 1765 AlaHisIleGlnPheLeuAenPheSerThrGluPro-----AenHisAspTYr 1780
 DB 3036 CTGAATATTCCTCCAGACATTTGAAGTGGAGGAGAACGCGACTGTGCTATGACTAT 3095
 QY 1781 ILGluIleArgAenGlyProTYrGluThrSerArgMetGlyArgPheSerGlySer 1800
 DB 3096 GTGAGGCTCTTGTATGGTCTTGATTCACAGCTGGGCTGGTGCATTCGTGATCC 3155
 QY 1801 GluLeuProSerSerLeuLeuSerThrSerHisGluThrValTYrPheHisSerAsp 1820
 DB 3156 GGGCCACCAAGAGATTTATTCATTCGATTCGATTCATTTATTCATTTCCACACTGAT 3215
 QY 1821 HisSerGlnAenArgProGlyPheLeuGluTYrGlnAla 1834
 DB 3216 GACACATCAACAAAGAGGATTCATATTAAGATACAAAGC 3257

RESULT 13

US-09-432-473-1
 ; Sequence 1, Application US/09432473
 ; Patent No. 6365715
 ; GENERAL INFORMATION:
 ; APPLICANT: ARLETH, ANTHONY J.
 ; APPLICANT: WILLETT, ROBERT N.
 ; APPLICANT: ELSHOURBAGY, NABIL A.
 ; APPLICANT: LI, XIAOTONG
 ; TITLE OR INVENTION: HUMAN CARDIAC/BRAIN TOLLUOID-LIKE PROTEIN
 ; FILE REFERENCE: ATG-50038-D1
 ; CURRENT APPLICATION NUMBER: US/09/432,473
 ; EARLIER FILING DATE: 1999-11-01
 ; EARLIER APPLICATION NUMBER: 08/991,408
 ; EARLIER FILING DATE: 1997-12-16
 ; EARLIER APPLICATION NUMBER: 60/034,471
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 5145
 ; TYPE: DNA
 ; ORGANISM: HOMO SAPIENS
 ; US-09-432-473-1

Alignment Scores:

Pred. No.: 5.1e-48 Length: 5145
 Score: 769.50 Matches: 245
 Percent Similarity: 37.83% Conservative: 131
 Best Local Similarity: 24.65% Mismatches: 331
 Query Match: 4.53% Indels: 287
 DB: 3 Gaps: 33

US-10-016-248-2 (1-1104) x US-09-432-473-1 (1-5145)

QY 879 AsnAenHisGluCysIleTYrSerIleGlnThrGlnProGlyLysGlyIleGlnLeuLys 898
 DB 1023 GATTAACCACTAATTCATTAAGAAAGAAACATCCAGCCAGGTCAA----- 1067
 QY 899 AlaArgAlaPheGluLeuSerGluGlyAspValLeuIleValTYrAspGlyAsnAenAen 918
 DB 1068 -----GAGTACAAATTTCTGAGATGGAGCCTGGAGAAAGCAAAAC 1106
 QY 919 SerAlaArgLeuLeuGlyVal-----PheSerHisSerGluMetMetGlyValThrLeuAen 937
 DB 1107 TCA-----CTTGAGAGAAAGATATGATTTGACAGATATGACATATGCGACAGAAC 1157
 QY 938 SerThrSerSerSerLeuTrpLeuAspPheIleThrAspAlaGluAenThrSerLysGly 957
 DB 1158 ACCTTCTCAAGGGGAGATGTTCTGATACCATCTTCCCTCCCGCT----- 1202
 QY 958 PheGluLeuHisPheSerSerPheGluLeuIleLysCysGluAspProGly---ThrPro 976

DB 1203 -----GATGATATATGCACTAGCTCT 1223
 QY 977 LysPheGlyTYrLysValHis---AspGluGlyHisPheAlaGlySerSerValSerPhe 995
 DB 1224 GCATTTGTGCAAGCAACCCGCTTAAGCAAAAGAGATATGCAAGCA----- 1271
 QY 996 SerCysAspProGlyTYrSerLeuArgLysSerGluLeuLeuLysSerGlyGlu 1015
 DB 1271 ----- 1271
 QY 1016 ArgArgThrTrpAspArgProLeuProThrCysValAlaGluCysGlyGlyThrValArg 1035
 DB 1272 -----AGAAAGCTGTATGATGTCCAGCA---TGCGAGAAACTCTACAA 1313
 QY 1036 GlyGluValSerGlyGlnValLeuSerProGlyTYrProAlaProTYrGluHisAenLeu 1055
 DB 1314 ---GAATCCAAATGGCAACCTTCCCTCCAGGATTTCCCAATGGCTTCTTACACA 1370
 QY 1056 AenCysIleTrpThrIleGluAlaGluAlaGlyCysThrIleGlyLeuHisPheLeuVal 1075
 DB 1371 CACTGATCTGAGAGCTTCTGTGACCCAGGGAGAAAGATTTGTTAAATTTTACAAAG 1430
 QY 1076 PheAspThrGluVal-----HisAspValLeuArgIleTrpAspGlyPro 1091
 DB 1431 ATGATCTATACAAAGAGTATTGTGCTGTGTATGACTATTAAGATTAAGACGGGTAC 1490
 QY 1092 ValGluSerGlyValLeuLeuLysGluLeuSerGlyProAlaLeuProLysAspLeuHis 1111
 DB 1491 TGGAGAAATCACTCTCTGTGTAGATTCTGTGGGACCAATTCCTGAACTTCTTACT 1550
 QY 1112 SerThrPheAsnSerValValLeuGlnPheSerThrAspPhePheThrSerLysGlnGly 1131
 DB 1551 TCTACAGACAGAGAAATGAGATTGATTT----- 1580
 QY 1132 PheAlaIleGlnPheSerValSerThrAlaThrSerCysAenAspProGlyIleProGln 1151
 DB 1580 ----- 1580
 QY 1152 AenGlySerArgSerGlyAspSerTrpGluAla---GlyAspSerThrValPheGlnCys 1170
 DB 1581 -----CGTAGAGCAGATTAATGGTAGAAAGGCTTTGACAGCTGTATGAA--- 1628
 QY 1171 AspProGlyTYrAlaLeuGlnGlySerAlaGluIleSerCysValIleIleGluAenArg 1190
 DB 1628 ----- 1628
 QY 1191 PhePheTrpGlnProSerProProThrCysIleAlaProCysGlyGlyAspLeuThrGly 1210
 DB 1629 -----GGATCTGTGAGAGTACGTAA 1655
 QY 1211 ProSerGlyValIleLeuSerProAenTYrProGluProTYrProProGlyLysGlyCys 1230
 DB 1656 AATGAAGACAGATTCAGTCTCCCAATTATCTGATGACTATGCCCAATGAAGAAAGT 1715
 QY 1231 AspTrpLysValThrValSerProAspTYrValIleAlaLeuValPheAsnIlePheAen 1250
 DB 1716 GTGTGAGAAATPAACGTGTGTGAGAGCTACACAGCTGGGGCTGACCTTTCAGTCTTGA 1775
 QY 1251 LeuGluPro-----GlyTYrAspPheLeuHisIleTYrAspGlyArgAspSer 1266
 DB 1776 AATGAAGACATGACAAATTTGCTTATGACTACCTGGAAGTATGAGATGAACCAAGTGA 1835
 QY 1267 LeuSerProLeuIleGlySerPheTYrGlySerGlnLeuProGlyArgIleGluSerSer 1286
 DB 1836 AATTAACCTTTGATGAGCGCTTCTGTGTATGACAAACCTGAAGACATTAAGACTTAC 1895
 QY 1287 SerAenSerLeuPheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPheVal 1306
 DB 1896 TCCATATCTTGTGATGAAGTGTCTTCTGACGGAACCTGGAACAAAGACGGTGTCT 1955
 QY 1307 IleAspTYrThrGluAsnProArgGluSerCysPheAspProGlySerIleLysAenGly 1326
 DB 1956 GCTACTTTTAAAGAG---GAAGATGAGGTGCAAACT----- 1994


```

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 648..3689
OTHER INFORMATION: /product= "human mt11 protein"
US-08-866-650-4

```

```

Alignment Scores:
Pred. NO.: 3.89e-48 Length: 3919
Score: 768.50 Matches: 245
Percent Similarity: 37.83% Conservative: 131
Best Local Similarity: 24.65% Mismatches: 331
Query Match: 4.52% Indels: 287
DB: Gaps: 33

```

US-10-016-248-2 (1-3104) x US-08-866-650-4 (1-3919)

```

QY 879 AsnAhnIeGluCysIleTySerIleGlnThrGlnProGlyIysGlyIleGlnLeuLys 898
DB 1419 GATACCACTTAATCATATACAGAAACATCCAGCCAGGTCA----- 1463
QY 899 AlaArgAlaPheGluLeuSerGluGlyAspValIleuLysValTyraSpGlyAAsnAAsn 918
DB 1464 -----GAGTACAAATTTCTGAAGATGAGCCCTGGAGAAAGTAAAC 1502
QY 919 SerAlaArgLeuLeuGlyVal---PheSerHisSerGluMetGlyValThrLeuAen 937
DB 1503 TCA-----CTTGAGAAAGATATGATTTGCACAGATACGACTATGCGCAGAAC 1553
QY 938 SerThrSerSerLeuThrProLeuAspPheIleThrAspAlaGluAenThrSerLysGly 957
DB 1554 ACCCTTCAGAGGGGAGATGTTTGGATACCATTTCTCCCTCCCT----- 1598
QY 958 PheGluLeuHisPheSerSerPheGluLeuLysCysGluAspProGly---ThrPro 976
DB 1599 -----GATGATATGCGCATACGTCCT 1619
QY 977 LysPheGlyTyTyLysValHis---AspGluGlyHisPheAlaGlySerSerValSerPhe 995
DB 1620 GCATTTGGTCAGGACCCGCTTAAGCAAGAGATATCCACAGCA----- 1667
QY 996 SerCysAspProGlyTyTySerLeuArgGlySerGluGluLeuLeuCysLeuSerGlyGlu 1015
DB 1667 ----- 1667
QY 1016 ArgArgThrTrpAspArgProLeuProThrCysValAlaGluCysGlyGlyThrValArg 1035
DB 1668 -----AGAAAGCTGTATAGATGCCAGCA---TGTGGAGAAACTTACAA 1709
QY 1036 GlyGluValSerGlyGluValLeuSerProGlyTyTyProAlaProGlyGluHisPheLeuVal 1055
DB 1710 ---GAATCCATAGCACTTCTCTCCAGGATTTCCCATGAGCTACCCCTTCTTACACA 1766
QY 1056 AsnCysIleThrThrIleGluAlaGluAlaGlyCysThrIleGlyLeuHisPheLeuVal 1075
DB 1767 CACTGCATCTCGAGAGTTTCTGTGACCCCGGAGAGAAAGATTTGTTTAAATTTTACACAG 1826
QY 1076 PheAspThrGluGluVal-----HisAspValIleuArgIleThrAspGlyPro 1091
DB 1827 ATGATCTATACAAAGAGTATTGTGCTGTATGACTATATGAAGTAAAGACGGGTAC 1886
QY 1092 ValGluSerGlyValLeuLeuLysGluLeuSerGlyProAlaLeuProLysAspLeuHis 1111
DB 1887 TGGAGAAATACACCTCTCTGTGTAGATTTCTGTGGAGACAAATGCTGTAAGTTCTTACT 1946
QY 1112 SerThrPheAsnSerValIleuGlnPheSerThrAspPheThrSerLysGlnGly 1131
DB 1947 TCTACAGACAGCAAGATGTGATTTGAGTTT----- 1976

```

```

QY 1132 PheAlaIleGlnPheSerValSerThrAlaThrSerCysAsnAspProGlyIleProGln 1151
DB 1976 ----- 1976
QY 1152 AsnGlySerArgSerGlyAspSerTrpGluAla---GlyAspSerThrValPheGlnCys 1170
DB 1977 -----CGTAGCAGCAGTAAATGGGTAGAGAAAGACCTTTCAGACTCTATGAA--- 2024
QY 1171 AspProGlyTyTyAlaLeuGlnIleGlySerAlaGluIleSerCysValLysIleGluAenArg 1190
DB 2024 ----- 2024
QY 1191 PhePheTrpGlnProSerProProThrCysIleAlaProCysGlyGlyAspLeuThrGly 1210
DB 2025 -----GGCATGTGTGGAGGTAGATACCTAA 2051
QY 1211 ProSerGlyValIleLeuSerProAenTyTyProGluProTyTyProProGlyLysGlyCys 1230
DB 2052 AATGAAAGACAGATTCAGTCTCCCAATATCTGTATGACTATCGCCCGGTGAAGAAATGT 2111
QY 1231 AspTrpLysValThrValSerProAspTyTyValIleAlaLeuValPheAsnIlePheAen 1250
DB 2112 GTGTGAAATAACAGTGTCTAGAGCTACACGTGGGCTAACCTTTCAGTCTTTGAG 2171
QY 1251 LeuGluPro-----GlyTyraSpPheLeuHisIleTyraSpGlyArgAspSer 1266
DB 2172 ATTGAAGAAGCATGACAAATGTGCTTATATGACTACCTGGAAGTAAAGATGAACAGTGA 2231
QY 1267 LeuSerProLeuIleGlySerPheTyTyGlySerGlnLeuProGlyArgIleGluSerSer 1286
DB 2232 AATAGCCCTTATAGGGGGTTCGTGTATGACAAACCTGAAGACATAAGATCTTACC 2291
QY 1287 SerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPheVal 1306
DB 2292 TCCAACTACTTTGGATGAAGTGTGTTCTGACGGAACGTGAACAAACAGAGTTTGTCT 2351
QY 1307 IleAspTyTyThrGluAsnProArgGluSerCysPheAspProGlySerIleLysAsnGly 1326
DB 2352 GCTAACTTTTAAAGAG---GAAGATGAGTGTGCCAAACCT----- 2390
QY 1327 ThrArgValGlySerAspLeuLysLeuGlySerSerValThrTyTyTy-----Cys 1343
DB 2391 GACCGTGAAGGCTGTGAGCAGGAGTGTGAACACTCGGCGCACTTACAGTGTGCTGT 2450
QY 1344 HisGlyGlyTyTyGluValGluGlyThrSerThrLeuSerCysIleLeuGlyProAspGly 1363
DB 2451 GAGCCTGGCTATGAG-----CTGGGCCAGAC--- 2477
QY 1364 LysProValIleTrpAsnAsnProArgProValCysThrAlaProCysGlyGlyGlnTyTyVal 1383
DB 2478 -----AGAGAGAGCTGTGAAGCTGTGTGTGGACTTCTTACC 2516
QY 1384 GlySerAspGlyValValLeuSerProAsnTyTyProGlnAsnTyTyThrSerGlyGlnIle 1403
DB 2517 AAACCTTAACGGGACCATTAACACCCCTGCTGGCCCAAGAGTACCTCTTAATAAGAAC 2576
QY 1404 CysLeuTyTyPheValThrValProLysAspTyTyValIlePheGlyGlnPheAlaPhe 1423
DB 2577 TGTGTGTGCAAGTGTGTGACCAACCAAGTACAGAAATTTCTGTGAAGTTTGAAGTTTGT 2636
QY 1424 HisThrAlaLeuAsnAspVal-----ValGluValHisAspGlyHisSer 1438
DB 2637 GAATTGAAAGGCAATGAAGTGTGCAAAATATGATTAATGTGAGATGTGAAGTGTCTTCC 2696
QY 1439 GlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThrGlyGlySerLeuProLeu 1458
DB 2697 TCTGAGCTAAACCTGACATGCAAAATCTGTGGCGGTGAAGTCCGTAAGTATC---ACA 2753
QY 1459 AlaThrSerAsnGlnValLeuIleLysPheSerAlaLysGlyLeuAlaProAlaArgGly 1478
DB 2754 TCCCACTTCAACAAATATGAGAAATTTCAAAATCTGCAATCTGTATCCAGAAAGGCT 2813
QY 1479 Phe-----HisPheValTyTyGlnAlaValProArgThrSerAlaThrGlnCysSerSer 1496

```

```

Db      2814  TTCAAGACATTTTTC-----TCAGCAAAAGATGATGCTCT--- 2852
Qy      1497  ValProgluProArgTyrGlyLysArgLeuGlySerAspPheSerValGlyAlaIleVal 1516
Db      2853  -----AAGGATAATGGTGGATGTCAGCAGCAATGTGTCAACAGCATGGGAGCTACATG 2906
Qy      1517  ArgPheGluCysAenSerGlyTyrAlaLeuGlnGlySer-----ProGluIle 1532
Db      2907  ---TGTCAATGCCCTAATGATTTGTGTCTACATGACATGAATAAACATGATTGCAAGAAAGCT 2963
Qy      1533  GluCysLeuProValProGlyAlaLeuAlaGlnTyrPheValSerAlaProThrCysVal 1552
Db      2964  GAGGT-----GACGAGAAATCCACAGCTCAAGT----- 2993
Qy      1553  ValProCysGlyGlyAenLeuThrGluArgArgGlyThrIleLeuSerProGlyPhePro 1572
Db      2994  -----GGCCCTCATCCACAGTCCCACTGGCCA 3020
Qy      1573  GluProTyrLeuAenSerLeuAenCysValTyrPylsIleValValProGluGlyAlaGly 1592
Db      3021  GACAAAGTACCCAGCAGAGAAAGAAATGCACTTGGGAAATCAGGCCACTCCGGCCACCGA 3080
Qy      1593  IleGlnIleGlnValValSerPheValThrGluGlnAen-----TyrAspSer 1608
Db      3081  ATCAATATGACCTTATGATGATTTGAGATTGAGCAGCATCAAGATGTCTTATGACAC 3140
Qy      1609  LeuGluValPheAspGlyAlaAspAenThrValThrMetLeuGlySerPheSerGlyThr 1628
Db      3141  TTAGAAGTATTTGATGGAGAAACAGAAAGTCACCGATTTGGACGACATGTGGGCAAC 3200
Qy      1629  ThrValProAlaLeuLeuAenSerTherAenGlnLeuTyrLeuHisPheTyrSerAsp 1648
Db      3201  AAGATACCGATCCCTTGTGGCTACGGAATAAAGTTGTGGTGTGTTCTGAT 3260
Qy      1649  IleSerValSerAlaAlaGlyPheHisLeuGluTyrLysThrValGlyLeuSerSerCys 1668
Db      3261  GCATCTGTTCAAAGAAAGGCTTTCAAGCCACACATTTACA----- 3302
Qy      1669  ProGluProAlaValProSerAenGlyValLysThrGlyLysArgTyrLeuValAsnAsp 1688
Db      3302  ----- 3302
Qy      1689  ValValSerPheGlnCysGluProGlyTyrAlaLeuGlnGlyHisAlaHisIleSerCys 1708
Db      3302  ----- 3302
Qy      1709  MetProGlyThrValArgArgTyrAsnTyrProProProLeuCysIleAlaGlnCysGly 1728
Db      3303  -----GAGTGTGGC 3311
Qy      1729  GlyThrVal-----GluGluMetGluGlyValIleLeuSerProGlyPheProGly--- 1745
Db      3312  GAGCAGATTGAAGCAGATCAAAACCAAGAGATGTGATCAGATGTCAGATTGGTGAT 3371
Qy      1746  ---AsnTyrProSerAenMetAspCysSerTyrPylsIleAlaLeuProValGlyPheGly 1764
Db      3372  AACAACTAACCGAGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3431
Qy      1765  AlaHisIleGlnPheLeuAenPheSerThrGluPro-----AsnHisAspTyr 1780
Db      3432  CTTGAATATATCTTCCAGACATTTCAAGTGAAGGAAAGCAGACTGCGCTATGACTAT 3491
Qy      1781  IleGluIleArgAenGlyProTyrGluThrSerArgMetGlyArgPheSerGlySer 1800
Db      3492  GTGAGCTCTTGTGATGCTTGTGATTCACAGCTGTGGGCTTGTGCTGATTCGTGGATCC 3551
Qy      1801  GluLeuProSerSerLeuLeuSerThrSerHisGluThrValTyrPheHisSerAsp 1820
Db      3552  GGGCCACCGAAGAGATTTATTCATTGGAATTCAGTTTAATTCATTTCACACATGAT 3611
Qy      1821  HisSerGlnAenArgProGlyPheLysLeuGluTyrGlnAla 1834

```

```

Db      3612  GACACATCAACAAGAGGATTTATATATAGATACAAAAGC 3653
RESULT 15
US-09-021-287-4
; Sequence 4, Application US/09021287
; Patent No. 5981717
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hoffman, Guy G
; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,287
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/866,650
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Benet J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296, 93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3919 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: Linear
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 648..3689
; OTHER INFORMATION: /product= "human mTII protein"
US-09-021-287-4
Alignment Scores:
Pred. No.: 3,896-48 Length: 3919
Score: 768.50 Matches: 245
Percent Similarity: 37.83% Conservative: 131
Best Local Similarity: 24.65% Mismatch: 331
Query Match: 4.52% Indels: 287
DB: 2 Gaps: 33
US-10-016-248-2 (1-3104) x US-09-021-287-4 (1-3919)
Qy      879  AsnAenHisGluCysIleTyrSerIleGlnThrGlnProGlyLysGlyIleGlnLeuLys 898
Db      1419  GATTAACACCACTAATCATTAAGAAAGAAACATCCAGCAGTCA----- 1463
Qy      899  AlaArgAlaPheGluLeuSerGluGlyAspValLeuLysValTyrAspGlyAenAsnAsn 918
Db      1464  -----GAGTACAAATTTCTGAAGATGAGAGCTGGAAGAAATTAAC 1502
Qy      919  SerAlaArgLeuLeuGlyVal-----PheSerHisSerGluMetMetGlyValThrLeuAsn 937

```

```
Db 1503 TCA-----CTTGAGAAAGATATGATTTGACAGATATCATGACTATGCCAGAAC 1553
Qy SerThrSerSerSerLeuThrLeuAspPheIleThrAspAlaGluAsnThrSerIleGly 957
Db 1554 ACCTTCTCAAGGGGAGGATGTTCTTGATACCAATTCCTCCCTCCCG- 1598
Qy PheGluLeuHisPheSerSerPheGluLeuIleGlySerGluAspProGly---ThrPro 976
Db 1599 -----GATGATATGAGCATACGCTCT 1619
Qy LysPheGlyIleGlyValHis---AspGluGlnHisPheAlaGlySerSerValSerPhe 995
Db 1620 GCATATGGTCAGAACCCGCTTAAGCAAGAGATATGCCACAGGCA----- 1667
Qy SerGlyAspProGlyIleGlySerLeuAlaGlySerGluLeuGlySerGlyGlu 1015
Db 1667 ----- 1667
Qy ArgArgThrTrpAspArgProLeuProThrCysValAlaGluCysGlyGlyIleThrValArg 1035
Db 1668 -----AGAAAGCTGTATGATGTCACGA---TGTGAGAAAGACTCTTCA 1709
Qy GlyIleValSerGlyIleValLeuSerProGlyIleProAlaProIleGluHisAsnLeu 1055
Db 1710 ---GAATGCAATGGCAACCTTCTCCACAGATTTCCCAATGGCTACCTTTTACACA 1766
Qy AsnCysIleTrpThrIleGluAlaGluAlaGlyCysThrIleGlyLeuHisPheLeuVal 1075
Db 1767 CACTGCATCTGAGAGAGTTTCTGTGACCCAGGGGAGAAAGATGTTTAAATTTTACACG 1826
Qy PheAspThrGluGluVal-----HisAspValLeuArgGlyIleTrpAspGlyPro 1091
Db 1827 ATGATCTATACAGAGATGTTGTGCTGCTGATACATATATGAAGTAAGACGGGTAC 1886
Qy ValGluSerGlyValLeuLeuGlyGluLeuSerGlyProAlaLeuProIleGlyAspLeuHis 1111
Db 1887 TGGAGAAATCACCCTCTCTGTGTAGATTTCTGTGGAGCAAAATTCCTGAAAGTCTTACT 1946
Qy SerThrPheAsnSerValLeuGlnPheSerThrAspPheThrSerIleGlyGlnGly 1131
Db 1947 TCTACAGACAGCAAAATGTGATGAGTT----- 1976
Qy PheAlaIleGlnPheSerValSerThrAlaThrSerCysAsnAspProGlyIleProGln 1151
Db 1976 ----- 1976
Qy AsnGlySerArgSerGlyAspSerTrpGluAla---GlyAspSerThrValPheGlnCys 1170
Db 1977 -----CGTAGCAGCAGTAAATTTGGTAGAAAGGCTTTGACAGCTGTATGAA--- 2024
Qy AspProGlyIleValLeuGlnGlySerAlaGluIleSerCysValLysIleGluAsnArg 1190
Db 2024 ----- 2024
Qy PhePheTrpGlnProSerProProThrCysIleAlaProCysGlyGlyAspLeuThrGly 1210
Db 2025 -----GCCGATCTGTGAGAGTGAGATACGTAA 2051
Qy ProSerGlyValIleLeuSerProAsnTrpProGluProIleProGlyIleGlyGlyCys 1230
Db 2052 AATGAGAGACAGATTCAGTCTCCCAATATCTCTGATGACTATGACCCGATGAAAGATGT 2111
Qy AspTrpIleValThrValSerProAspTrpValIleAlaLeuValPheAsnIlePheAsn 1250
Db 2112 GTGTGAAATAACAGTGTCTGAGAGCTACACGTGGGGCTGACCTTTAGTCTTTGAG 2171
Qy LeuGluPro-----GlyTrpAspPheLeuHisIleGlyAspGlyArgAspSer 1266
Db 2172 ATTGAAGACATGACATTTGCTTATGACTACCTGAGAGTGAAGATGAAACAGTGA 2231
Qy LeuSerProLeuIleGlySerPheTrpGlySerGlnLeuProGlyIleGlySerSer 1286
Db 2232 AATAGCCCTTTGATGGCGCTTCTGTGTTATGACAAACCTGAGACATTAAGATCTACC 2291

Qy 1287 SerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPheVal 1306
Db 2292 TCCAAATACCTTGTGATGAAGATTTGTTCTGACGGAACTGTGAACAAACAGAGGTTTGT 2351
Qy IleAspTrpThrGluAsnProArgGluSerCysPheAspProGlySerIleLysAsnGly 1326
Db 2352 GCTAATCTTTTAAAG---GAAAGTAGTGTCGCAACCT----- 2390
Qy ThrArgValIleSerAspLeuLysLeuGlySerSerValThrTrpTrp-----Cys 1343
Db 2391 GACCGTAGAGGCTGTAGAGAGGAGTGTGAAACACTCTGGAGACTACCAAGTGTGCTT 2450
Qy HisGlyGlyIleGlyValGluGlyThrSerThrLeuSerCysIleLeuGlyProAspGly 1363
Db 2451 GAGCTGGCTATGAG-----CTGGGCCCAAGC--- 2477
Qy LysProValTrpAsnAsnProArgProValCysThrAlaProCysGlyIleGlyIleThrVal 1383
Db 2478 -----AGAAAGAGCTGTGAAGCTGCTGTGTGAGCTTCTTAC 2516
Qy GlySerAspGlyValValLeuSerProAsnTrpProGlnAsnTrpThrSerGlyGlnIle 1403
Db 2517 AAACCTTAACGGCACCATACCAACCCGTGGCTGGCCCAAGAGACTACCTCTAATAAGAAC 2576
Qy CysLeuTrpPheValThrValProLysAspTrpValPheGlyGlnPheAlaPhePhe 1423
Db 2577 TGTGTGTGGCAAGTGTGTGCACCAACCAAGTACAAATTTCTGTGAAGTTGAGTTTTT 2636
Qy HisThrAlaLeuAsnAspVal-----ValGluValHisAspGlyHisSer 1438
Db 2637 GAATTCGAAGGCAATGAAGTTTGCAAAATATGATTTGTGAGATGTGAGAGTGTCTTTC 2696
Qy GlnHisSerArgLeuLeuSerSerLeuSerGlySerHisIleGlyGlySerLeuProLeu 1458
Db 2697 TCTGAGCTTAACCTCATGCGCAAAATTCGTGGCGCTGAAGTCCGAAAGTGATC---ACA 2753
Qy AlaThrSerAsnGlnValLeuIlePhePheSerAlaLysGlyLeuAlaProAlaArgGly 1478
Db 2754 TCCAGCTTCACAAATATGAGATTAATTCAAATCTGACAACTACTGTATCCAAAGAGGCG 2813
Qy Phe-----HisPheValTrpGlnAlaValProArgThrSerAlaThrGlnCysSerSer 1496
Db 2814 TTCAAAGACATTTTTC-----TCAGCAAAAGATGAATGCTCT--- 2852
Qy ValProGluProArgTrpGlyLysArgLeuGlySerAspPheSerValGlyAlaIleVal 1516
Db 2853 -----AAGATTAATGTGATGTCAGACAGCAATGTGTCAACAGATGGAGGATACATG 2906
Qy ArgPheGluCysAsnSerGlyTrpAlaLeuGlnGlySer-----ProGluIle 1532
Db 2907 ---TGTCAATGCCGTATGATGATTTGTGCTACATGCAATAAATGATTTGCAAGAGACT 2963
Qy GlnCysLeuProValProGlyAlaLeuAlaGlnTrpAsnValSerAlaProThrCysVal 1552
Db 2964 GAGTGT-----GAACGAAGATCCACAGTCCAAAGT----- 2993
Qy ValProCysGlyAsnLeuThrGluArgArgGlyThrIleLeuSerProGlyPhePro 1572
Db 2994 -----GGCTCATACACAGTCCCACTGGGCA 3020
Qy GluProTrpLeuAsnSerLeuAsnCysValTrpLysIleValValProGluGlyAlaGly 1592
Db 3021 GACAAGTCCCAACAGCAAGAAATGCACTTTGGAAATTCAGCCGCACTCCGGGACACGA 3080
Qy IleGlnIleGlnValIleSerPheValThrGluGlnAsn-----TrpAspSer 1608
Db 3081 ATCAAAATAGCTTTAGATTTGAGATTTGAGAGCATGACAGATCAAGATGTGCTTATGACAC 3140
Qy LeuGluValPheAspGlyAlaAspAsnThrValThrIleLeuGlySerPheSerGlyThr 1628
Db 3141 TTAGAAGATTTGATGAGAAAGCAAGAAAGTCAACCGATCTTTGAGACATATGTGGCAAC 3200
```

```
QY 1629 ThrValProAlaLeuLeuAsnSerThrSerAsnGlnLeuTyrLeuHisPheTyrSerAsp 1648
Db 3201 AAGATACCAATCCCTGCTGCTACAGAAATGAATGTTGCGTTGTTGTTGCTGAT 3260
QY 1649 IleSerValSerAlaIleGlyPheHisLeuGluTyrIleThrValGlyLeuSerSerCys 1668
Db 3261 GCATCTGTCAAGAAAAGGCTTTCAGCCACACATTCCTACA----- 3302
QY 1669 ProGluProAlaValProSerAsnGlyValIleThrGlyGluArgTyrLeuValAsnAsp 1688
Db 3302 ----- 3302
QY 1689 ValValSerPheGlnCysGluProGlyTyrAlaLeuGlnGlyHisAlaHisIleSerCys 1708
Db 3302 ----- 3302
QY 1709 MetProGlyThrValArgArgTyrAsnTyrProProProLeuCysIleAlaGlnCysGly 1728
Db 3303 -----GAGTGTGGC 3311
QY 1729 GlyThrVal-----GluGluMetGluGlyValIleLeuSerProGlyPheProGly--- 1745
Db 3312 GGACGATTGAAGCAAGAAATCAAAACCAAGAGATCTGACTCATGCTCACTGCTTGTGAT 3371
QY 1746 ---AsnTyrProSerAsnMetAspCysSerTyrPlySileAlaLeuProValGlyPheGly 1764
Db 3372 AACAACTAACCCAGAGACAGCTGACTGTAATGCTATTAGTATCAGAACGGGCTCTCGA 3431
QY 1765 AlaHisIleGlnPheLeuAsnPheSerThrGluPro-----AsnHisAspTyr 1780
Db 3432 CTTGAATTATCTCTCCAGACATTTGAAGTGGAGAGACAGACAGCTGTGCTATGACTAT 3491
QY 1781 IleGluIleArgAsnGlyProTyrGluThrSerArgMetMetGlyArgPheSerGlySer 1800
Db 3492 GTGAGCTCTTGTAGGCTTGTGATTCACACAGCTGTGGGCTTGGTCATTCGTGATCC 3551
QY 1801 GluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTyrPheHisSerAsp 1820
Db 3552 GGGCCACCAAGAGATTATTCATTGAGATTCAGTTTAATTCATTTCACACATGAT 3611
QY 1821 HisSerGlnAsnArgProGlyPheLysLeuGluTyrGlnAla 1834
Db 3612 GACACAATCAACAAGAGGATTCATATAGATACAAAAGC 3653
```

Search completed: October 18, 2004, 16:04:05
Job time : 1208 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 10:04:21 Search time 33 Seconds
(without alignments)
6237.911 Million cell updates/sec

Title: US-10-016-248-2

Perfect score: 16985

Sequence: 1 MACAPPALLPCLSLSDCC.....RSGPVDPSTLPQSHRSPK 3104

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A COMB pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B COMB pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A COMB pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B COMB pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS COMB pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1 pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1850.5	10.9	3594	US-09-911-842A-4	Sequence 4, Appli
2	1794	10.6	3571	US-09-911-842A-2	Sequence 2, Appli
3	1555	9.2	2489	US-09-911-842A-5	Sequence 5, Appli
4	1390.5	8.2	3623	US-09-341-461-2	Sequence 2, Appli
5	1375	8.1	1947	US-09-612-314A-52	Sequence 52, Appli
6	1362	8.0	2039	US-09-612-314A-52	Sequence 52, Appli
7	1362	8.0	2039	US-09-612-314A-52	Sequence 52, Appli
8	1267.5	7.5	1847	US-09-285-385C-10	Sequence 10, Appli
9	1267.5	7.5	1847	US-09-285-385C-10	Sequence 10, Appli
10	1133	6.7	1466	US-09-285-385C-4	Sequence 4, Appli
11	1133	6.7	1466	US-09-285-385C-4	Sequence 4, Appli
12	1133	6.7	1537	US-09-285-385C-4	Sequence 4, Appli
13	1133	6.7	1537	US-09-285-385C-4	Sequence 4, Appli
14	771	4.5	1012	US-09-991-408-2	Sequence 2, Appli
15	769.5	4.5	1013	US-09-432-473-2	Sequence 2, Appli
16	769.5	4.5	1013	US-09-432-473-2	Sequence 2, Appli
17	768.5	4.5	1013	US-09-432-473-2	Sequence 2, Appli
18	768.5	4.5	1013	US-09-432-473-2	Sequence 2, Appli
19	768.5	4.5	1013	US-09-432-473-2	Sequence 2, Appli
20	768.5	4.5	1013	US-09-432-473-2	Sequence 2, Appli
21	760.5	4.5	1023	US-10-140-002-200	Sequence 20, Appli
22	750	4.4	1015	US-09-285-385C-2	Sequence 2, Appli
23	746	4.4	986	US-09-285-385C-19	Sequence 19, Appli
24	744.5	4.4	788	US-08-572-225-1	Sequence 1, Appli
25	738.5	4.3	1013	US-08-866-650-3	Sequence 3, Appli
26	738.5	4.3	1013	US-09-021-287-3	Sequence 3, Appli
27	738.5	4.3	1013	US-09-240-473-3	Sequence 3, Appli

28	703	4.1	910	US-10-140-002-112	Sequence 112, App
29	702	4.1	986	US-08-872-757-4	Sequence 4, Appli
30	702	4.1	986	US-08-872-757-4	Sequence 4, Appli
31	666.5	3.9	830	US-08-110-158-4	Sequence 4, Appli
32	663	3.9	591	US-08-991-408-4	Sequence 4, Appli
33	663	3.9	591	US-08-991-408-4	Sequence 4, Appli
34	648.5	3.8	830	US-09-432-473-4	Sequence 4, Appli
35	630	3.7	830	PCT-US91-05059-2	Sequence 2, Appli
36	609.5	3.6	574	US-09-475-460A-32	Sequence 32, Appli
37	528	3.1	610	US-08-365-470-3	Sequence 3, Appli
38	528	3.1	610	US-09-209-668-19	Sequence 19, Appli
39	528	3.1	610	US-09-009-490A-89	Sequence 89, Appli
40	526	3.1	610	US-09-009-490A-89	Sequence 89, Appli
41	502	3.0	611	US-09-475-460A-32	Sequence 32, Appli
42	502	3.0	611	US-09-475-460A-32	Sequence 32, Appli
43	501	2.9	577	US-08-435-149-3	Sequence 3, Appli
44	485	2.9	730	US-08-872-757-2	Sequence 2, Appli
45	485	2.9	730	US-09-850-048A-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-911-842A-4
; Sequence 4, Application US/09911842A
; Patent No. 6656707
; GENERAL INFORMATION:
; APPLICANT: Amgen Inc.
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017/37592
; CURRENT APPLICATION NUMBER: US/09/911,842A
; CURRENT FILING DATE: 2004-07-24
; PRIOR APPLICATION NUMBER: US 60/222,438
; PRIOR FILING DATE: 2000-08-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 3594
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1757)..()
; OTHER INFORMATION: Xaa = any or unknown amino acid
US-09-911-842A-4
Query Match 10.9%; Score 1850.5; DB 4; Length 3594;
Best Local Similarity 21.3%; Pred. No. 8.4e-134;
Matches 789; Conservative 378; Mismatches 1170; Indels 1369; Gaps 149;
99 TCNEGYDLOGSKKTCMKVSDMPAMSDHPRVCRARCDALHPSGIIITSPNFIQYDN 158
493 TCNEGYDLOGSKKTCMKVSDMPAMSDHPRVCRARCDALHPSGIIITSPNFIQYDN 158
159 NAWCWITLANSKVKLAFFERDLEKGLDITLVGGGDDGQKTVLVNSQANSDSPH 218
540 -----PS-----CGKQPA 547
219 TPG-----SRIP-----SMGDIWROKTV-----LEICRDISSDARSVYRS 259
548 RPKMTQCLSGRGVYLSGVAVCAATSG-----KMSAKVGTAVCKD-----EA 591
260 PRTSNVAVLAPGTETIEQSCGDPGIPAYRRREGSRPHGDTLKECPAPFE-----LVG 314
592 PQLSCPDIDIAKTCGEQDSANATWQVPT-----AKNSGEKYSVHVPFTPPYLPYLP 645
315 QKAITCQKNQWAKKGCVPF-----CFPNTSPSGV-VLSPNPEYDGNH 360
646 DVATITVATTS-IGNQASCTFYIKVIDEPPVIDWC-----RSPPIQVKEKHPASWD-- 698
361 LHCWMLIARPSRIHLAFNDIDVEPQDFELVIKDGATAPVLGTFPS-IGNQPSSTSS 419
```

```

Db 699 -----EPQF-----SDNSGAELVITSHTGDMFPHGTVV 729
QY 420 GHVAREFQTDHSTGKRGFNITFTFRHNCBPDGVVNGKRKRDSDLSQLGSSISFLCDEG 479
Db 730 WYRA-----TDPSSNNNTCDI-HIVIKGSPCEVPEFTVNGDFICADDSAGVNSLSCKEG 783
QY 480 FLGTQGS-ETITCVLKEG-----SVVNSAVLRCEAPCG-----G 513
Db 784 YDFTEGPEKYCAFEDEGIMRPPYSTEMPCAIKRFANHGFKSFEMLYKTRCDMDLFX 843
QY 514 HLTPSSTIISPCMPGYKDA-----LSCAMVTEAOPGP-----I 549
Db 844 KFAFAFTTIGANNVPSFCNDADDIDCRLEBDLTKKYCIEVNYNENGFAIGCGWGANRL 903
QY 550 KITPDRKTEVNDLTLEVRDGR-----TYSAPLIGVYHG-----Q 585
Db 904 DYSDHDLVYQELFPDVGAKSRIRKTYPLSDPKQLFINITASVPLPEBNDLTELE 963
QY 586 VPQFLIST-----SNLYLFTSDKSHSDIGFOLRYETITLQSDHC-----LDGPIPV 633
Db 964 NOQRILKTELTITNRLKSTLNKEPMYS--FOLASEIVADNSLSETEKAFLCRPGSVL 1020
QY 634 NGQR-----HGNDFYVALVFCDSGYTLSDGEPECE-----PNFQMSPALPSCBA 681
Db 1021 RGMVCVCPGLGTSYSLHSTCEBCLMGSYODEGQLECKLCPRTHABYLHSHSVSECKA 1080
QY 682 LCGGFIQSSSG-----TI-----694
Db 1081 QCGQGTSSSGLETCEBCLGTYQPERGSRCLCEPTETTTVKRAVIDIACGVPCPVGB 1140
QY 695 -----LSGFRP--DEVPNN-----LNCTW-----711
Db 1141 FSRSGLTPCYPCRDYQPNAGKSFCLACPFYGTITITGATSTIDCSSFSSTPSAABESI 1200
QY 712 --IIEHSHGKVF--FTFHTFHL-----ESGHDYLLITENGSTQPLRLQTLG- 754
Db 1201 VPLVABGHSONKKEVSSQVFNHECFLLPNCCHNSGTCCQLGRGYVCLCPRG-----YTL 1252
QY 755 -----SRLPAPISAGL--YGNFTAQVRF-----IS 777
Db 1253 KCETDIDECSLPC-LNGGICRDQVGFTECELSGSGQI CEHINIECTISSPCLANKGTCT 1311
QY 778 DEEMSY-----BGFNTFSBYDLEPCBEBEVPAYISIRKGLQFGVDTLTJFSC-----FPG 827
Db 1312 DGLASRYCTCVKGMVGHCTDVNECOSBCLNNAVCQDVGG-----FSCKCPGFLG 1365
QY 828 YLEGRYA-----RITCLGRRRLMSSPLPRVACGNSVGTGOTGLSPNPNVAY 877
Db 1366 TRCEKXVDECLSQPCQNGATCKDGANSF--RC--QCPAGFTGTGCEL-----NI 1410
QY 878 N-----NNHECI-----YSIOTOPGKGIQLKARAFELSEGDVLKVDGNNSARLIG 924
Db 1411 NECQSPNCNQATCVDELNSYSCKQPG--FSHGRCETBQ-----SGNLDREVSG 1460
QY 925 VBSHSEMKV--TLNSTSSLMV--DFTT-----DAENT-----SKGPELHPS 963
Db 1461 IYGYVLLDVLPLHLAVTCAFMKSSDVINYGTPTISYALEDDKDNFTLTDYGVWLYVN 1520
QY 964 SFELI-----KCEDPGT-----975
Db 1521 GKEKINCPBVNDGIMHIAITWTSIGAMRYIIDELSDGGTGLSIGAKITGGGALVIG 1580
QY 976 -----PKF-----978
Db 1581 QEDDKKGEFNPAESFVGSISQNLMDVYLSPOQVYLLASSCEBELSRGNVLAMPDFLSG 1640
QY 979 ---GYVHDEGHA-----GSSVSPSCDPGYSLRSEBELLCIS 1013
Db 1641 ITGKYVVDSSMFCSDPSELBGSVPLHLPASGNRKSGKYSLFCDPGFQMVGNPNVQYCLN 1700
QY 1014 GERRTVDRPLPCV-AECG---GYVAGEVSGVLSGYPARVEHNNNCMTWTEAAGCTI 1069
Db 1701 --QGGWTOGFLPHCIRICGLPALENGFSABDFHAGSTVYQ-----CTS 1744

```

```

QY 1070 GLHFLVPTBEVHDVLIRINDGVEBSCVLLKEIS-----1102
Db 1745 G-YLLGDSRMFCXDNGSWNGISPSCLDVECAVSDCEHASCLNTNGSVYCSNPPYT 1803
QY 1103 -----GPALPKDLSTFNSVVLQSTDFEFTSKOGFAIOFS-----VST 1140
Db 1804 GGGKXCAEPVKKCAENPENGRS-----SGEITYT--VGTATVFSDEBHELGVST 1852
QY 1141 AT-----SCNDPGI PONGSRSGDSWEAGDSTVFOCDPGVALQSGAETS 1183
Db 1853 ITCLETEGWDRLRPCEAISCVPVPPEVGVDSGAFTYGSKYVYRCDKGYTLISGDEESA 1912
QY 1184 CVKIENRFPWQSPPTC-IAPCGDLTGSVYILSPNVEPFPKRECKOMKTVBPDYI 1242
Db 1913 CLASGS--WSHSPVCGLVKC-----SQEDINNK-----YIL 1944
QY 1243 A-----LVFNIFNIEBPGYDFLHIYDGRDLSPLIGSFYGSQULPGRITESNSLFLAFRSDA 1298
Db 1945 SGLTYLSTIASYSCENGYSL-----QGSULLECTASSGW-----DR 1979
QY 1299 SVSNAGFYIDYENPRESCPDGSIKNGTRVSGDLKLSVTTYHGHGVEGISTLSCT 1358
Db 1980 APPSCOLV-----SCGEPPIVKDAVITGNSNFTFGTVAVYCKEGYTLAGPDTIIC- 2029
QY 1359 LGPDGKPVNNRPVCTAPCGGQYVSDGVLSBPVPONYISGQICLFFVTYPKD-YVVF 1417
Db 2030 -QANGK--MNSNHOCIA-----VSCDEPVPVDHAS-----PETAHRLF 2065
QY 1418 GQFAFFHALNDVVEVHDGHSQHSRLSLSGSHGESLPLATSNQVLIKFSAKGLAPAR 1477
Db 2066 GDTAFYYCA-----DQYS-----LADNSQLCNQGMVNPVPPG 2098
QY 1478 GRHFYQAVPRTSATQCSVPEPRYG--KRLGSPFSGALYRPECNSGYALQSPBEIC 1534
Db 2099 -----QAVPRCIAFCEKPEPSVYSILESVSKAFAGSVVSFKCMGPFVLTNTAKIEC 2152
QY 1535 LFPVGAALQMVNSAPTCVPCGQNLTERGTLISGPFEPYLSLNCWVKIIVPBGAGIQ 1594
Db 2153 LR-----GGENS-----PSPL--SVQC-----2167
QY 1595 IQVVSFTEQWDSLEVEDGADNTVTLGSEFGTTPALNLSTSNQVLIHFYSDISVAA 1654
Db 2168 -----IPV-----2170
QY 1655 GRHLEKYTVGLSSCEPRAVPNSGVTKGERYLVNDVYSQCEBRYALQGHANISCMPTVR 1714
Db 2171 -----RCGEPPSIANGYPSGTNYSFGAVVAVYASHKGFYIKGEKSTC--EATG 2216
QY 1715 RWNYPPLCIAQCGGTVEEMEGVILSPGFPCNYPSPNMOCSWKIALPVGFAHIQFLNFT 1774
Db 2217 QMSKRTPTC-----HP--VSCNBPKEVNGFLHTTORTRES 2251
QY 1775 EPNHDYIIRNGPYETSRMGRFSGSELPSLLSTSHETTVYPHSDHSONRPPFLXEYA 1834
Db 2252 EARFOC--NPGYKA-----AGS--PVFVQANR--HMSHDAPLCTPL-----2288
QY 1835 YELQCPDEPFPANGIVRAGAVNYGQSVTEFCLGQYOLGHPVLTCQIGTRMND-HPLP 1893
Db 2289 ---NGKRPPIQNGFLKGBSEFVGSKVQFVCNGBEYELVGNMWTQ--KSGKMSKKRSP 2342
QY 1894 KCEVP--CGGNITSNGTVSPGFPSPSSQDCWMLITVYIGHVRLNLILQTEPSDG 1951
Db 2343 KC-VPTKCAEPPLBNQVLKE-----LASBEVMTISCKEGHALQ-GPSVLKCLPBGQ 2394
QY 1952 FITTWDEPOQTAPRLGVFTSRMAKTVQSSNQVLLKFRDAATGGLFALAFSAVPLTK- 2010
Db 2395 ---WNG-----SFPICKM 2404
QY 2011 ---CPEPTILNAAVYTNENEFNIGDIYRCLBFTVLVGNELCKLGTYIQFEGRPIC 2068
Db 2405 VLCSPPLIPFG-VPASSGALHFGSTVTKYLCVDGDFLGRGSLTILCOADS--TWSSPLPEC 2461

```

QY 2069 -EVAHPTNELTSTGVILSSQSYPGSFQFOTCSMLVREPDYINSLTVEFPLESEKQYDE 2127
 Db 2462 VPEECPOPEILN--GIHVOGLAVLSTLYTC-----KPEFL-----VGN 2501
 QY 2128 FEIIFDPSGO-----SPLLKASGNYSAPLIVTSSNSVULWMSDHAYNKGFIRYSAP 2183
 Db 2502 ATTLGNGQWLGKPKCKPIE----- 2523
 QY 2184 YCSLPRAPLHGLIGOTSTPGGSIHFCNAGYRLVHSHMAICTRHQGYHLSAIPLC 2243
 Db 2524 -CEPKEILNG-OFSSVSYQYQITTYFCNDRGFLBPKSLTCLJETD-----WDMDPSC 2577
 QY 2244 QALSCLGPBARKGMVFGKEYTGTAKVASCSEGYHQAQEAETACLTGLTGLMSNRVWP 2303
 Db 2578 DAHCSPPRIENGFEAGADRYGAMITISCFQVULGHAMQT--CEESG-MSSSS--P 2632
 QY 2304 QCVPVTC--PDV--SSISVEHGRWELIFE----- 2328
 Db 2633 TCVPIDGCLPRHIDFGDCTKVRDQGHFDQEDDMMEVRYLAHPQHLBATAKALENTKESP 2692
 QY 2329 ----TOYQFOAQLMIDPBYTYTGORVIRCOANGKSLDSTPTCTIISGELPT-PPN 2383
 Db 2693 ASHASHPLVGTWVSYSCEPVEYELLGIPVLICQEDGTWN--GTAPSCISIEC-DLPVAPEN 2749
 QY 2384 GHRIGTLVYGATAIFSCNSGYTLVGSVRBECMANGLMSGSEVRCLAGHGTEPEIYNH 2443
 Db 2750 GFLHPTOTMGSAAOYSCRGHILBESHLCLCNKMGSGTVPRCALISCKRPLNLNGS 2809
 QY 2444 INGENYSYRGSVVYQCNAGFRLIGMSVYRICQODHMSGKTPFCVPTICGHPGNVNLQ 2503
 Db 2810 IKDDVSYLVLYVECDGYILNGSKKRTQENRDMGHEPWCIPVDCGSPVPPTNGRVK 2869
 QY 2504 GNPFLNDVYKFCNPNBYMAEGARSCCLASGQSDMLPTCRILINCTDPHQENSVRQVH 2563
 Db 2870 GEEYTPQKELTYSCEBGFILBEGARSRLCLTNNGSYSGATPSCMPRCAPQOVNPGV--- 2925
 QY 2564 ASGPHRPSFGTYSYRCNMGFVLLGTPLVLSQCGDGTDRPRQCLAVSCGHPSPSPHQM 2623
 Db 2926 ADGLD-YGFKGEVAFHCLBEGVULQABRLTCQSNIGTDAEVPCKRATCCPPADLPQGF 2984
 QY 2624 SGGSYTVGAVVRYSCIGKRTLNGSTNMCGLDHMTGSLPHC-----SGT 2668
 Db 2985 NGSPFYHGHIOYQCFYGLKHNPSRRCJLPGNSWSSSPSCJLPCRCSTPLIOQGTINAT 3044
 QY 2669 SVGV-----CGD-PGLAHGIRLCSF 2689
 Db 3045 DLGCGKTVOJECRKFGLGLSEITCDANGOWSDVPLCEHAQCGPPLTIP-NAIVLEGLS 3103
 QY 2690 DPGTVNRFSCEAGHVLKGSSEBRTQANGSWSGQPECGVISCGNPGTPNSA----- 2740
 Db 3104 SEDNVVYISCRPGYTMQSSDLICTEKAIVSOPRYPTCEPLSCGPPTVAANAVALTEAHY 3163
 QY 2741 -----RVVF--SDGLVFS 2751
 Db 3164 ESKVYKRLCEGVWMDSDTDTFTCOQDGHWPBRTTSCPKKCPVPSMNTIRIRFGDDPQVN 3223
 QY 2752 SSVIVBREGYVATGLISRHCSVNGTGT--SDPECLVINCDBGIRANGRLGANDPRYN 2809
 Db 3224 ROVSVCABEFTHEGVWMTSCOPGTWEPFSDSCIPVYCGHPESPAHSGSVGNKHSFG 3283
 QY 2810 KATYQCVPGVYMHSHSVLSCTKDRTNMGTKVYCALMCKPRLPINKVAGSDPMWG 2869
 Db 3284 STIVYQDDPGYKLEGNHRI--COENQOMSGEVAVNCENRCETPAEPKNAVALENTTSG 3341
 QY 2870 SSVYVACLEGYOLSLPAVFTCEGNGSWTGLPOCFPVCGDPGV-PSRGRREDRGFSYRS 2928
 Db 3342 PSLIFSCHRGYTLEGSPEAHCTANGTWNHLTPCLKRCPVPRVPIRPNNAVLSEKEFVVDQ 3401
 QY 2929 SVSPSCRPVLYVGSPPRFQSDGTMGSGTOPSCIDPLTLTTCADBPVQFQIGNNSQ--Y 2986
 Db 3402 NVSIKCEBGLKNGNGVITCSPDETWHTNARC--EKISGSPSHVENAI--ARGVY 3455
 QY 2987 QVGSYVLFRCQKGVLLQGSTTRTCLPVLTWSGTRPDCVPHNCSQREPRTHANVGAIDLPS 3446

Db 3456 QYGDMTTYSYCSYMEGSLRSVCLENGTM--TPSPVCRAVCFP-----CQNGSVQCRPN 3509
 QY 3047 -----NGYTLITPARRASPSRVAPSTAPARMAAGASRPSAW 3084
 Db 3510 ACSCPDGMWGRIC-----EFPICILPLCLNGRCVAPYQDCDCTPW 3549
 RESULT 2
 US-09-911-842A-2
 ; Sequence 2, Application US/09911842A
 ; Patent No. 6656707
 ; GENERAL INFORMATION:
 ; APPLICANT: Amgen Inc.
 ; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF
 ; FILE REFERENCE: 0101/37592
 ; CURRENT APPLICATION NUMBER: US/09/911,842A
 ; PRIOR FILING DATE: 2001-07-24
 ; PRIOR APPLICATION NUMBER: US 60/222,438
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 3571
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-911-842A-2
 Query Match 10.6%; Score 1794; DB 4; Length 3571;
 Best Local Similarity 21.3%; Pred. No. 2,1e-129;
 Matches 799; Conservative 399; Mismatches 111; Indels 1354; Gaps 155;
 QY 74 CDPPIPERG---KALGSDFLGSSVQFTCNBYDLOGSKRTTCKVKSMPFAASDHP 129
 Db 378 CPALKPEBNGYFIQNTCNHFNACGVR--CHPGFVLVGSIIILCP---NGLWGSSES 431
 QY 130 VCARACDADLHGRPS--GIITSRNPFIQYDNNACWIMITALNPSKYIKLAFEEFDERGY 188
 Db 432 YCVRVTC-PLHRPKKHGHSICSTREMLYKTT--CL-----VACDEGRHLEGS 475
 QY 189 DTLTV-GDGGQDDQKTVLYMSQNASD-----SPHTPG----- 221
 Db 476 DKLTQGNQWQDPEBRPV---ERHCSTFGMPKDIIISPNCGKQPAKFGTICVSCROG 532
 QY 222 ---SRIPESM---SGDIMRQKTV---LEICDDISSDARSVSXKSPSTNAVELVAP 271
 Db 533 FILSGVEMLRCTTSG---KNNVGVQAAVCKDV-----EAPQINCPKIDEAK 576
 QY 272 GTEIEGSGCDPGIPAYGRREGSRFHHGDTLKFEQPAFE---LVGOKAITCQKNQW 326
 Db 577 TLEQDQSANVTWQIPY---AKDNGSEKVSVNHAFAPLPYLPFGDVAVIYTTATD-L 629
 QY 327 SAKKPGCVSCEPNTSPSGVVLSPVYREDYGNHLHCWMLILARPSRIHLAFNDIDVER 386
 Db 630 SGNQASCIPIHIXIDAEPPYIDCRSPR-----VQVEBKVAAASMD--EP 673
 QY 387 QPFLVYIKGATLAEAVLGTFS--GNOLPSITSSGHVARELFTDSTGKRGNNITFTTF 445
 Db 674 QF-----SDNSGAEIVITTSHTQDLPFQGETIVQYTA-----TDSGNNRKTDI-HIV 722
 QY 446 RHNECPDPGVYVNGKRFQGSLOGSSISFLCDEGFLGTGS--ETICVULEGSVVNSAV 504
 Db 723 KGSPPCEIPPTPVAGDPTCTPDNTGVNCTLTCLGYPFTBSTDKYCAVEDG--VMKPY 780
 QY 505 LRCEAPCGHLLTSPSG-----TILSPGMGFYKDA--- 534
 Db 781 TTMPROCAKRRFANHGFKSEMFYKARODDITLMKKFSAPFTTLGKVPSCSAEDI 840
 QY 535 -----LSCAVIEAOPGYP-----IKTFDRFKTEVNYDLEVRDGR- 571
 Db 841 DCRLEBNLTAKYCLEVNYVYENGFAIGCGWGAANRLDYSDFDLTVDETATISNASK 900
 QY 572 -----TYSAPLIGVYHGT-----QVQPLISTSNLYLFLSTD 604

```

Db 901 SRIKSAPLSDYKIKLIFENITASVPLPBERNDTLEMENQRLDITETITKLTANKD 960
Qy 605 KSHSDIGQLRYETI-----TLQSDHCLDPSI PVNGORHGN-----DFYVALVTFSQ 652
Db 961 PMYS---FQLASELLIDNSLSLETAKSPRCPSGVSILGRMCVNCPLGTY--NLBHFTC 1015
Qy 653 DS-----GYLSDGEPLCE-----PWFQSRALPCEALCS-----684
Db 1016 ESHCIGSQDEEGQ-LECKLCPGMYTEYIHSRNISOCKAQCKQGYVSYLETCECPL 1074
Qy 685 GFIO---GSSGTILSP-----GPP-----DFYNN 706
Db 1075 GTYQPKGSRCLSCPESTSTVKRGAVNISACVPCBEKFSRGLMPCPCRDYQPN 1134
Qy 707 -----LNTWITISGKVFFTFPHLESGHDYLLITENGSFTOPLRQLTGSRIPAP 760
Db 1135 AGKAFCLACP---YGTTPFAGSR-----ITECSFSSSTFSAABESVVP-P 1177
Qy 761 ISAG-----LYGNFTAQV---RFISDFMSYGFNITESEYDLEP 797
Db 1178 ASLGHKRRHISQVHECFENPCNSGTCQQLGRGVCLCPGLTGLK---CETDIDE 1234
Qy 798 CEEPEVAYISIRKGLQGVGDTLTSCFPGR---LBSTA- 834
Db 1235 CS---PLPCLMNGVCHDLVGEFICEPSGYTGQRCEBINECSSPCLNKGI CVDGVAG 1290
Qy 835 -RITCLG-----RRRLMSP-----LPRCY 854
Db 1291 YRCTCVGFPVGHCTEVENECQSPCLNNAVCEDOVGLCKCPGFLGTRCKNDECL 1350
Qy 855 AE-CGNSVTGQGT---LSPNF---PVNYN-----NNHECI---YGIOTQ 890
Db 1351 SOPCKNGATCKDGNAPRCLCAAGFTGSHCELINECQSPCNQATCVDELNSYCKQ 1410
Qy 891 PG-----KGIOKARAFELS-----905
Db 1411 PGFSGKRCTEGSTGFELD---FEVSGIYGYMLDGLPSLHALTCTFWKSSDDMNYGT 1467
Qy 906 -----EGDYL-----KYV 913
Db 1468 PISYAVDNGSDNTLLLDYNGWVLYVNGREKINCSVANDGRWHHAIITWTSANGIKWY 1527
Qy 914 -----DGN-----NNSARLLGV-----FHSBMMGVTLNSTS-SLWLDPI 948
Db 1528 IDKLSDBGAGLSVGLPIPGGALVVGQEQDKKEGFEFSPESP---VGSISQNLNM-DYV 1583
Qy 949 TD-----AEMTSKGFELHFSF-----ELIKCEDPRTKFG-----Y 980
Db 1584 LSPQVKSIALTSCPEELSKGNVLAMPDFLSGIVGKVKIDSKSIFCSD--CPRLGGSVPHL 1641
Qy 981 KYHDEGHFAGSSVVSFSCDPRYSLRGSEELLICLSGERRTWRPLPTVACGCTVAGVSG 1040
Db 1642 RTASEDLKSGSKVNLPCDPGFQVGNPVQCLN--QGQWTOPLPHC-----1685
Qy 1041 QVLSPEYRPAEYENLNCIWTIEACCTIGL---HLVPTDEYVHDVLRITWDGPVBSG 1095
Db 1686 ERISCVPPLENGFHS--ADDFYAGSTVYQCNNGYVILLGDSRMFCTDNGSMNGVSPSC 1743
Qy 1096 VLLKELS-----GPRALPKDLASTN 1115
Db 1744 LDVDECAVSGDCSEHASCLNDGSIYCSVPPRTGDKNCAPIKAKABGFENGHSS-G 1802
Qy 1116 SVYLQSTPFTSKOGFAIQ-----FSVSTATSCNDPQIPONGSRGDS 1159
Db 1803 ELYTVGAETVFSQOEYQVLMGVTKITLESSEWNHLIPYKAVASGKPAIPENGCIEELA 1862
Qy 1160 WEAGDSTVFOCDPGYALOGSAEISCVKLENRFWQSPPTCTAPCGGLITGSSVILSPN 1219
Db 1863 FTFGSAVYTRCKNGKGYTLADGKSSCLANS---WHSHPVC-----1900
Qy 1220 YEPYPPGKCDMWKTVSPDYIALVFNFLEPGVDFLHYDGRDLSPLIGSF---YG 1276

```

```

Db 1901 ---EPV-----KCS-----SPE-----NINNG---KYILSLTYLSTASYCDTQYS 1936
Qy 1277 SOLPRIESSNSLFLAFRSDASVSNAGFVIDYTEMPRESCFDPGSIKNGTRVGSDDLKG 1336
Db 1937 LQPSILIECTAGSIW-----DAPRACHLVF-----CCEPAIKAVITGNPFR 1982
Qy 1337 SSYTYCHAGGYVEGTSTLSCLGPDGKPVNNPVPCTAPCGGQVSGDGVVLSNPYQ 1396
Db 1983 NVTYTYCKRGYTLAGDITIECL--ADGK--WGRSDQCLIA-----VSCDEBP 2025
Qy 1397 ---NYTSGQICLYFVVPKDYVVGQFAPFHTLNVVEVHDSQHSRLSSLSGHTBE 1454
Db 2026 IVDHASPERA-----HRLFQDIAFYCS-----DGYT-----2052
Qy 1455 SLPLATSNQVLIKFSAKGLAPARGFHFVQAVPRTSATOCSSVPEPRYV---KRLGDSFS 1511
Db 2053 ---LADNSQLLCNAQKQWVPEG-----QDMPRCIAHFCEKRPVSIVSILESVAKARA 2103
Qy 1512 VGAIVRECNQVYALQSPSEIECLEVPBALAOMNVSAPTCVVPCCGNLTERGTTLSPGF 1571
Db 2104 AGSVVSFKMBGFVINTAKIECM---RQGMN-----2133
Qy 1572 PEPYINSINCWKIIVPBGAGIQVVSFVTEQUNDSEVFDGADNTVTLGSPSGTTP 1631
Db 2134 PEPM---SIQC-----IP 2143
Qy 1632 ALLNSTNQVLYHFYSDISVSAAGHLEYKTVGLSSCEPAVPSNGVTGERLYVNDVYS 1691
Db 2144 V-----RCGEPPSINMGVAGSNVSPGAMVA 2169
Qy 1692 FQCEPVYALOGHAHISCMPTVRMNYPRPLCINQCGTVEMBEVILSPGPGYPSNM 1751
Db 2170 VSCNKFYIKKEKSTC--EATQWSPILPTC-----2199
Qy 1752 DCSWKIALPVGGAHIOFLNTEBPNHDYIEIRNGPYETSRMMGRFSGSELPSLSTSH 1811
Db 2200 ---HPVSCG-----BPP---KYENGFLE-----H 2217
Qy 1812 ETTYVPHSD-HSNRPGPK-LEYQAYELQ-----ECDDPEPANGIYRG 1853
Db 2218 TTRGFEESEVRVQCPGKYSVSPVQANRHHSESPLMCVPLDCKPPIQNGFMKG 2277
Qy 1854 AGYVNGQSVFTECLPGVQLTGHPVLTQHGTRNRMHDPL-PKC-----EYPCGN-----1902
Db 2278 ENFEVGSVQVFCNMGVYELVDSWTCQ--KSGMKNSKNCKMPAKCPPEPLLENQVL 2335
Qy 1903 ---ITSSNGTVYSPGPSPSYSSQDCWMLTIVIGHGVRLNLSLQTEBSGDFITWDGPQ 1960
Db 2336 KELTTEVGVV---TFSCKE-----GH-VLOGPSVLKCLPSQC---MN---2370
Qy 1961 QDAPRLGVFTRSMKKTVOSSNQVLLKFHNDATGGIFALAFSAIPLTK---CPPTIL 2017
Db 2371 ---DSFVCKIVLTCPPELI 2387
Qy 2018 PNAEVTENEENFNGIDIVRYRCLPFTLVGNEILTCKL-GTYLOFESPPPIC-EVHCEPTN 2075
Db 2288 -SFGVPIPSALHFGSVYKXSCVGGFRLRGSTLLQCPDGTW---SSPLPECVVEECPQ 2443
Qy 2076 ELLDSTGTVILSGYPSGYPQFQTCMWLVREVPYINISLVEYFLSKQYDEFEI FDGPS 2135
Db 2444 EETPN--GILDVQSLAYLSTALYTC-----KPGFEL-----V 2473
Qy 2136 GQSPDLKALSGNYAPLITVSSNSVYLRWSSDAVYRKGFKIRYSAFYCSLPPRAPLHG 2195
Db 2474 GNTTLTGENGHW-----LOGKPTCKAIECLKREILNG- 2507
Qy 2196 ILGQTSQPGGSIHFGCNAGYRLVGHSMALCTRHPOQYHLSBAIPLCOALSGLPEAPK 2255
Db 2508 KFSYTDLHYQVTVYVSCNRGFRLEGPALTCLETGD---WDVADAPCSNAIHQDSPRIE 2563
Qy 2256 NGWVFGKXYVTKKAVYSCSEGYHLOAGATATACLDITGMSRNNVPQCVPTC-----2310
Db 2564 NGFVEGADYSYGAIIITYSCPPGF--QVAGHAMQTECSG-WSSS--IPTCMPIDCGLRPH 2618

```

```

Qy 2311 ---PDVSIIVH-----RWLIFE-----TOYQFQ 2334
      |:::|:::|
Db 2619 IDBGDCKLADDOGYFEQEDDMMEVPVYTHPPHLCANAKTMENTESPATSSNPLYG 2678
      |:::|:::|
Qy 2335 AOLMLICDPGYXYTGAVIRCOANGKWSLSDPTPCRIISGCELPPIPNGRIGTLISVG 2394
      |:::|:::|
Db 2679 TMTSYTCNPGYELLGNPVLLIQEDGTWN--GSAPSCISIEDDLTAENGLARTETSMG 2736
      |:::|:::|
Qy 2395 ATAFSCNSGYTLVGSVRBEAMANGLSGSEVRLACHCTPEPTVNGHNGENYSTVRS 2454
      |:::|:::|
Db 2737 SAVQYSCPGHILVGSDLRLCLENRKWSGASPRCEAISCKKPNVNMNGSIXGSVYTLST 2796
      |:::|:::|
Qy 2455 VVYQCNAGPRLIGSVIRICODDHMSGKTPCEVPIITGCHGNPNVNGITQGNQFLANVVK 2514
      |:::|:::|
Db 2797 LYIECDGIVLNGERTTCODDKMDEDEPLCTIPVDCSPVSNAGVGRGDEYTFQKEIE 2856
      |:::|:::|
Qy 2515 FVCPNGYMAEGAASQCLASGOWSDMLPTCRIINCTDPGHOENSVROVHSGPHRFSFGT 2574
      |:::|:::|
Db 2857 YTCNEGFLLEGARSRVCLANGSWGATPDCVPRCAPPPOLANGVTE----GLDYGFPMK 2911
      |:::|:::|
Qy 2575 TVSYRCHGTYLLCTPVLSGCGDGTWDRPPRQCLVSCGHPGSPHSQMGSDYTVGAVV 2634
      |:::|:::|
Db 2912 EVTFHCHGTYLLHGAAPLTCOSDGNMDAIEIPLCKPVNCGPREDLAHGFPNGFSFIHGHI 2971
      |:::|:::|
Qy 2635 YSCIGKRTLVGNSTRMC-----2652
      |:::|:::|
Db 2972 QYOCFPGYKLGHSNRCLNSGWSGSSPCLPRCSTPVIEGTVNGTDFDCKAARIQ 3031
      |:::|:::|
Qy 2653 -----GL-----DGHWTGSLPHCSGTSVGVCD-PGIPAHGIRLSDGPDPTVMR 2696
      |:::|:::|
Db 3032 CFKGFKLGLSEITCEADGQMSGSPHCEHTS---CGLPMIPAHFISSTSMWE-NVIT 3087
      |:::|:::|
Qy 2697 FSCGAGVNLGSSERTCOANGWSGSOPEGCVISCGNPETPSNA-----2740
      |:::|:::|
Db 3088 YSCRSYVLIQSSDLICTEKQWMSQPYVCEPILSGSPSVANAVATGEATYESEYKLR 3147
      |:::|:::|
Qy 2741 -----RVFSDGLVFSSIYEC 2758
      |:::|:::|
Db 3148 CLEGYMTDITDTTCTCKDGRMFERPERSCSPPKCPLEBNTITHLVHGDDFSVMQVSVSC 3207
      |:::|:::|
Qy 2759 REGYVATGLSRHCSVNGTWTG--SDPECLVINGDGPRIANGRLGNDFRYNKTVTYQC 2816
      |:::|:::|
Db 3208 AEGTTFEGVNIISVQLDGTWEPSPDSGSPVSCGKESPEHGVSGKTFESTIITYQC 3267
      |:::|:::|
Qy 2817 VPGVMESHRVSVLSTKRTWNGTKPVCAKLMCKPPLIPNGKVSGSDFEMGSSVTYAC 2876
      |:::|:::|
Db 3268 EPGYELEGNERERV--COENQWMSGVAICKEETRCETLEFLNGKADIENRTTGPNVYVSC 3325
      |:::|:::|
Qy 2877 LEGYQLSLPAVFTCEGNGSWTGLPQCFPVPCDPCV-PSRGREDGEGFYRSVSFSCH 2935
      |:::|:::|
Db 3326 NRGISLGSFSEBAHTEGTWHPVPLCKPNCVPFVYIPENALISEFEYVDONVSTKCR 3385
      |:::|:::|
Qy 2936 PPLVLVSPRRFCQSDGTWSGTQPSCTIDPLTTCADGVPQFGIQQNSQGVQVSTVLF 2995
      |:::|:::|
Db 3386 EGFLOGHGIIITCNPDETWOTSAK---EKISGCPRAHYENALARGVH-YQYDMITYS 3441
      |:::|:::|
Qy 2996 CQKGYLLQSSSTTRCTCLPNLTWSTGTPDCVPHHCRQPEPTPHANYGALDLS-----M 3047
      |:::|:::|
Db 3442 CYSGYMTEGFLRSVCLNGTWT--SPPIC-RAVCRFP---CQNGGICQRPNACSCPEGM 3495
      |:::|:::|
Qy 3048 GYTLITPARASPRVAPSTAPARMAAGQASRPSAW 3084
      |:::|:::|
Db 3496 GRUC-----BEPICTILPCLNGCVAPYQCCDPPGM 3526
      |:::|:::|

```

```

; FILE REFERENCE: 01017/37592
; CURRENT APPLICATION NUMBER: US/09/911,842A
; CURRENT FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: US 60/222,438
; PRIOR FILING DATE: 2000-08-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 2489
; TYPE: PR
; ORGANISM: Homo sapiens
; US-09-911-842A-5

Query Match          9.2%; Score 1555; DB 4; Length 2489;
Best Local Similarity 20.9%; Pred. No. 4.9e-111; Indels 1224; Gaps 142;
Matches 679; Conservative 332; Mismatches 1021;

Qy 58 KTSVLTVQVGVSGQHNNCPDGPGEIPERK-RUGSDPRLGSSVQVTCNEGYDLOGSKRITCMK 116
      |:::|:::|
Db 88 KNSVMTGAKDRCKRKSGRNPPDPVNGMVHVIKIGIQSGQIKYSCGTGYRLIGSSSATCI- 146
      |:::|:::|
Qy 117 VSDMFAASDHRVPCRAMCDALHRSBGIIITSPNPIQYDNNAHCWMIIT-ALNPSKVI 175
      |:::|:::|
Db 147 ISGDTVIMDNETPICDRIIPC-----GLPPIITGDFISTNRENFHGSVVTYRCNPGSGG 201
      |:::|:::|
Qy 176 KLAPEERDLERGYDTLVGSDGQDQKTVLVMSQNAACSHPHTPSGRIPESMGDIWRQ 235
      |:::|:::|
Db 202 KRYFE-----LVGEPS-----LYCTSN-----DDOYG-IW-- 225
      |:::|:::|
Qy 236 KMTVLIEICRDISSDARSQSVKSPKTSNAVELVAGTEIEQSGCDP---GIPAYGR 291
      |:::|:::|
Db 226 -----SGRAPQCIIP-----NKPPIPVNENGILVSDNR 253
      |:::|:::|
Qy 292 EGRFPHHGDLTKECOPAFELVQKAITCQKNQMSAKKGCYFSCFPNTSPSGVTLSP 351
      |:::|:::|
Db 254 --SLFSLNEVEERCOFGFVMPKPRRYKCOALKMKEBELPSCSRVC---QPPEDVLHAE 307
      |:::|:::|
Qy 352 NYPEDYGNHLHCWLLIARPEIRIHLAFNDIDVEPODPFLVINDGATAEAPVLGTSGNQ 411
      |:::|:::|
Db 308 RTQRDKDNF-----SFGQEVFS-----CEPGD---LNGAISMRCSTPGQDMS-- 347
      |:::|:::|
Qy 412 LPESITSSGHVAREFOTDSTKRGFNITFTFRHNECPD-PGVPNNGK-RFGDSLQLG 469
      |:::|:::|
Db 348 -PAA-----PCVYSQDPMFGQLNGVLRPVNLQLG 379
      |:::|:::|
Qy 470 SSISFLDSEGLTQSGSETITCVLKSGSVVWNSAVLRCEAPCGHLTSPSGTLLSPGMPG 529
      |:::|:::|
Db 380 AKYDFVCDEGF-OLKGSASAYCVLAGMESLMSNSVPVCE---QIFCPSPVPIPNG-- 430
      |:::|:::|
Qy 530 FYMDALSCAMVIAQGYPIKIFDRKTEVND-----TLEVR 568
      |:::|:::|
Db 431 -----RHTEKPLEVF--PRGKAVNTICDPHDPRTGSPDLIGESTRICTSDPQ 475
      |:::|:::|
Qy 569 DGRTVS--APLIGVYHGTVQPOPLISTSNLYLLFTGDKSHD--IGFQURVE----- 617
      |:::|:::|
Db 476 GNGWMSPARRCIILGHQAP-----DHLFPAKLTQTNASFPICSTLSKYCREPEYVG 529
      |:::|:::|
Qy 618 ---TTLQSDHCLD-----GPIVNGQRHG-NDPYGALVTFSCDSG 655
      |:::|:::|
Db 530 RPSPIIT-----CLDNLWSSPKDVCKRKSCTPPDPVNGMVHVIITDQVGSRIYNSCTTG 584
      |:::|:::|
Qy 656 YTLSDPELECEBN---FQMSRALPSECL-CCGFIQSSGTTLSPPFPFYNNNLNCTW 711
      |:::|:::|
Db 585 HRLIGHSSAACILISGNAHWSYTPICQRIIPC----- 617
      |:::|:::|
Qy 712 ILETSHGKGVFTFHTFHLESCHDYLLITENGSTFQPLRLQLTGSRRLPAPISAGLYGNFTA 771
      |:::|:::|
Db 618 -----LPPTIANG----- 625
      |:::|:::|
Qy 772 QVRPISDF-SMSYEGFNITSEYDLBCEBPVAVYSIRKGLQFGVGDITLTFSCFPG-- 827
      |:::|:::|
Db 626 -----DFISTNRENFH-----YGSVVTYRCNPGSGG 651
      |:::|:::|

```

```

RESULT 3
US-09-911-842A-5
; Sequence 5, Application US/09911842A
; Patent No. 6656707
; GENERAL INFORMATION:
; APPLICANT: Amgen Inc.
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF

```

QY 828 ---YRLBGTARITCLIGRRR--LWSSPLPRCAECGNSVTGTGTLSPN--PPVNNNN 880
 Db 652 RKVEIUGBPSIYCTSDNDQVIMSGAPCIT-----PKNCTPPNENG 696
 QY 881 HECTIYSIQTOPGKGIQLKARAFELSEGDVLKVYDGNNSARLLGVFSEMMGVTLNSTS 940
 Db 697 -----LIVSDNRS-----LFLINEV----- 711
 QY 941 SSLWLDFTIDAMENTSKEGELHFSSEFELICEDEPCTPRFGYKADHDEGHFASSSVSFSCDPG 1000
 Db 712 -----VEFRQPG 719
 QY 1001 YSLGSEBELLCSGERRTWRPLPTCVACGCGTVRGEVSGQLSPGPAPYEHNLNCTWT 1060
 Db 720 FVMKGPBRVCOALNK--WEPELPSGSRVC-----QPP----- 751
 QY 1061 IEABAGCTTGLHFLVFDTEBVHDLRIWDGVBESVILLKELSGPALPKDLHSTFNSVILQ 1120
 Db 752 -----DVLH----- 755
 QY 1121 FSTDFTSKQGFATQFSVSTATSCNDPGIPONGSRGSDSWEAGDSTVFOCDPGYALQSGA 1180
 Db 756 -----ABRTQRDXDNFSFGQEVYFSCPEYDILRGAA 786
 QY 1181 EISCVKIENRFPMQSPPTCIAPCGDLTGP--SGVILSPNYPEPYPRGKCDMKVTVP 1238
 Db 787 SMRCTPGD--WSPAPTCEVXSCDDFMGOLLNGRLFP----- 823
 QY 1239 DYVIALVFNLEPGDPLHIYDGRDLSPLIGSFQSOLPGRISSNSLFLA----- 1293
 Db 824 -----VNIQLGAKVDVCD-----EGFOLKG-----SSASYCVIAGMEST 858
 QY 1294 FRSDASVSNAGFVITYENPRESCFDPGSIKNGTRVGSDLKL--GSSVYVYGH----- 1344
 Db 859 WNSSVPCBOIF-----CPSPVPIPNGRHKGPLEVFPFGKTYVITCDPDPDNG 907
 QY 1345 GGYEBESTSLSCILGPDGKPVNNPPRYCTABCGQYVSGDGVLSPNYQNTYSQJIC 1404
 Db 908 TSEDLIGESTRICTSPDQNGVWSSPAPRC-----GIL----- 940
 QY 1405 LYEVTVKQVYVFGQAFHFTALNDVVEVHDGHSQ--HSLLSLSLGSHTGESLPIATSN 1462
 Db 941 -----GHCOAPDHFLEFKLKTQTNASDPFGTIS- 968
 QY 1463 QVLIKESAKGLAPARG-----HFYQAVPR--TSATQSSVPEBRYGK-RLGSDFSVG 1513
 Db 969 ---LKECREPREYGRPSITCLDNLWSS--PKDVCRKRSKCTPRPDVNGMVAHYITDIQV 1024
 QY 1514 AIVRECNAGYALOGSBEIECLFVPGALAQMNVSAPTCV--VPCGNLTERRTILSPGP 1572
 Db 1025 SRINYSCTTGHRLIGHSSAECI--LSGNAHMSKTPICORIPGCLPPTIANGDPISTNRE 1083
 QY 1573 EPLNLSNCWAKIVVEGAGIQIOWVSFVTEQNMWDLLEVFDGADNVTMLGSFSGTIVA 1632
 Db 1084 NFHYGS-----VVTYRCNPGSGRKYVE-----LVGERS----- 1112
 QY 1633 LINTSNOULYLFHYSDISVASAGFHLKYTVGLSSCEP--AVPS-----NGVKTB 1682
 Db 1113 -LYCTSDN-----DQVGIWSPAPQCIIPNKNCTPPNENGLVSD 1151
 QY 1683 R---YLVNDVVSFOCEPGYALOGHAHISCMPTGVRWNTYPPPLCIAQCGGTVBEMGVIL 1739
 Db 1152 NRSLEFLNEVERRCOPFVMKGPBRVCO--ALNMBEPELPSGSRVC----- 1197
 QY 1740 SPGFPGNYPNNDCWKIALPVGFAGHIOFLNSTEPNHDYLEIRNGVETSRMMGRFSG 1799
 Db 1198 -----QPPDVL----- 1204
 QY 1800 SELPSSLSTSHETTVYFHSDSQNRPGKLEVOAYELOECDDPEPFANGIVRAGAYNG 1859
 Db 1205 -----HAERTQ-----DKONF-----SPG 1219
 QY 1860 QSVTFECLRGYOLTHPVLTCQHGTVNRMDHPLPKCEVPCGNITSS--NGTVISBPFGPS 1917

Db 1220 QEVFYSCEBGYDLRGASNRCT--PDGWSPPAPFCEVXSCDDFMGOLLNGRLFP----- 1273
 QY 1918 PYSSQDCVWLITVPIGHGVRLNLSLQTEPESGDEFTITWDGPGQOTAPRLGVFTRSMAKKT 1977
 Db 1274 -----VNIQGAUV-----DIVCD-EGFQJMG----- 1294
 QY 1978 VOSSNOVLKFKHRDAATGIFALIAFSAVPLTK-----CPPEITLPNAEVVTEN--EERNIGD 2033
 Db 1295 --SSASYCVL-----AGMESLWSSVPECOJFCFSPPIVPINGRHTGKLEVPFEGK 1344
 QY 2034 IYRACLP-----GFTLVGNEILTCXL-----GYLOEBSPPICEV--HCPTBELILD 2080
 Db 1345 ANNYTCDHPDNGTSEFDLIGESTRICTSDPGONGW--SSPAPCGILGHCOA----- 1395
 QY 2081 STGVILSOSYPGSPQFOTCSWLVEBDYINSLTVEYFLSEKQYDEFEIPDPSGOSPL 2140
 Db 1396 -----PD-----HFLFAKLKTQTNASDPFGTSLX 1420
 QY 2141 LKALSGNSAPLITVSSNSVYLKWS--DHAYNRKGFIRISAPYCSLPAPRLHGFILQ 2199
 Db 1421 YECREYYGRPSTICLNLV--WSSPKDYCKRS-----CKTPPPVNGMVHT 1468
 QY 2200 TSTQGGSIHFGCNAGYRLVGHSMACITRHPOGYLMEAPPLCOALSCGLPEAPKNGV 2259
 Db 1469 TDIQVSRINYSCTTGHRLIGHSSAECILSGNTAH--WSTKTPICORIFCGLPPTIANGDF 1527
 QY 2260 FG---KEYVTGKAVYSCSEGVH---LOAGAETAECL---DTLWS-----NR 2299
 Db 1528 ISTNRENFHSGVVTYRCNLGSRGRKVELGEPISYCTSDNDQVIMSGAPQCIIINX 1587
 QY 2300 NVPP-----OCVP-----YTC-----PDVSSIS--- 2317
 Db 1588 CTPPVENGILVSDNRSI--SLANEVEFRQDQFVWKGRRAVCOALNMBELPSCSRYC 1647
 QY 2318 -----VEHGRWRLIFETOYQOAOQLMLICDEGYUYTGORVIRCOAMKMSIGDSTPTCRI 2372
 Db 1648 QEPPEILHGEHTPSHQDNFSPQGEVYFSCBEGYDLRGASLHCTPOGWS--PEAPRCAY 1705
 QY 2373 ISC---GELPFPNGHRICTLV-YGATAIFSCNSGYTLVGSRYRBCMANG--LWSGS 2424
 Db 1706 KSCDDFLQOL--PRGRVLPFLNLQLAGKAVFVDEGFRLGSSVSHCVLWGMSSLYMNS 1762
 QY 2425 EYRCLAGHCGTPEPVPVNGINGE---NYSYGSVYQCN-----AGRLIGMSYRICOQ 2475
 Db 1763 VPVCEHITCEPRLALNGRHTGTSPGDIPLYKEISITDPPDPDGMFTNLIGESTRICTS 1822
 QY 2476 DHR---WSGKTPFC-VBITCGH-----PGNPVNLTOGNQFNLDVVKFVGNP 2519
 Db 1823 DPHGNGVWSSPAPRCSELSVRAGHCKTPEQFPFASPTIPIINF-----EPVGTSLNYEGRP 1878
 QY 2520 GYMAEGARSGOCLASGQMSDMLPTCRILINCTDPGHQENSVQVNASGHRHSFCTTYSYR 2579
 Db 1879 GYFGMFSIS--CLNLVWSSVEDNCRKRSKCPPEPFGMVAHINTD---TOFSTVWYS 1933
 QY 2580 CNHGFYLLGTVLSQCGDG--TMDRPPQCLVASCGHGPGSPHPSQMSGDSYT----- 2629
 Db 1934 CNEGRFLIGSSTTCLVSGANNVTWDKAPICEIIS--EPPTISNGDYYSNNRISFH 1989
 QY 2630 VGAVVRYSC-----IGKRLVGNSTRMCGLDGHWGSLPHCSGTSVGVCDPG 2677
 Db 1990 NGTVVTVQCHTGPDEQLFELVGERSIY--CTSKDQDVGVWSSPPRCISITN--KCTAPE 2045
 QY 2678 IPAIGIRLGDS---FDGTVWRFCEGNYLARGSEKTCQANGWSSQPECGVYISGNP 2734
 Db 2046 V-ENAIRPGRNRSFSLTEIYFRQDQFVWVSHVTVQCTNGWGPRLPHCSRV--QOP 2103
 QY 2735 GTPSNARVFS--DGLVSSSIYECREGYATILSRHSVNGTWTGSPBECVINGD 2792
 Db 2104 PEILHGEHTLHQNFPGQGEVYFSCBPSYDLRBAALHCTPOGMSPEAPRCTVXSCDD 2163
 QY 2793 --PGIPANGRLGNDPFFYNKTVTYQCVGYMMESHVVS--VLSCXTKRTWNGTVPVCKAL 2848

```

Db      2164 FLAGLPHGRVLLPLNLQLGAKVSFVCEGFRLLKGRSASHCVLAGMK-ALMNSVYVCEQI 2222
Qy      2849 MCKRPPLIPNGKVGS---DEMSSSVTYAC-----LEGYLSLPAVTCENG 2894
Db      2223 FCPNPAILNGHRTGTPGDIPEYKEISYACDTHPRDRMTFNLIG-BSSIRCTSDROGNG 2281
Qy      2895 SMTGELPQC---FVFCGDPGVPSRGRREDGFSY---RSVSFSCHPPLVTVGSPRPF 2948
Db      2282 VMSSPAPRCELVSVAACDPDPKIQNGHIGHVSLYLBGMTISYICDPBGVLVKGFLFC 2341
Qy      2949 QSDGTWGTGTOPSCIDPLTTTCADGVPQPGIQNN---SQGYOVGSYVLFRCCQKGYLLQGS 3005
Db      2342 TDQGWGQLDHYCKE---VNCSPF-LFMNGISKELEMKVYHGDYVTLKCEDGYTLGGS 2397
Qy      3006 TTRTCLPRLTWSTGTPP 3021
Db      2398 PWSQCQADRRWD--PP 2411

RESULT 4
US-09-341-461-2
; Sequence 2, Application US/09341461
; Patent No. 6586389
; GENERAL INFORMATION:
; APPLICANT: Hammond, Timothy G.
; APPLICANT: Verroust, Pierre J.
; TITLE OF INVENTION: Cubilin Protein, DNA Sequences Encoding Cubilin
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: D6148
; CURRENT APPLICATION NUMBER: US/09/341,461
; CURRENT FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: PCT/US99/01259
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 40
; SEQ ID NO 2
; LENGTH: 3623
; TYPE: PRT
; ORGANISM: rat
; FEATURE:
; OTHER INFORMATION: amino acid sequence of rat cubilin protein
US-09-341-461-2

Query Match      8.2%; Score 1390.5; DB 4; Length 3623;
Best Local Similarity 20.7%; Pred. No. 6.3e-98;
Matches 702; Conservative 369; Mismatches 1069; Indels 1251; Gaps 130;

Qy      101 NEGYDLOGSKRITCMKVSDFMAVSDHRRPVCRAWCDALHGPSGIIT--SPNFPLOYDN 158
Db      443 NGGTICIDGINGFTC---DCTSSWTGVYCGTPQAACGILSGTGTAFYAGSPN--DYYIG 496
Qy      159 NAHCVMITLALNPKVYIKLAEFPDL-----RGYDTLVGDCGQDGDQKTVLYMSQNA 212
Db      497 NVNCFWIVRT--DEEKVLGVTFTFPDLSEASNCPRXY--LQIGDSDSADPFLGRY----- 548
Qy      213 CSDSPHPHPSRIPEMSGDIRQKMTVLEICRDISSDARSGSRKSKTNSIAEVLAPG 272
Db      549 C-----GSRPPGIC-----SSANALYF----- 566
Qy      273 TEIEGSCGDPGIPAYGRREGSRFHGDTLKFECQAPFELVGOKAITCQKNNQSAKKPG 332
Db      567 -----GLYSEVYRSGRF-----TARWEALKE 589
Qy      333 CVESCFNFTSPSGVVLSPNYPEDYGNHLHCVMILARPBSRIHLAFNDIVDP-----QF 388
Db      590 CGGILITDNY-----GSITSPGYGNYPPGRDCVMQVLVNPNSLITFTFGTSLSESGNCSK 645
Qy      389 DELVIKDGATAEAPVLGTFPSGNQLPSSITSSGHVABLEFOTDHSNGKRGFNTTTPRH 448
Db      646 DYLEIRDPFGQDPLVGLKFTCSLSTPLKTTGPARLIGFGSDSETSDKGFITVLTIOSD 705
Qy      449 -EC-----DPPGVVNGKR-----FGDS 465
Db      706 LDCGANTDTDTGELLPLPLSPFGSGRCVYVLTIOAQGEQIVINFTGVELESQMGCGTY 765

```

```

Qy      466 LQGSSTISFL---CDESGLTQSGEITTCVLKSGSVY-----NSAVLRG--BAPC 511
Db      766 IEVGDGSLRLKLC-----GNETLPPIRSVNKKWIRLRDALYOKASFRADYVAC 817
Qy      512 GGHLTSPSGTILSPGMPGFYKALASCAMVLEAOPGYIKITFDREK---TEVAYDTLEV 567
Db      818 GGMARG-EGFRSRSFYNAAPGRRTCRMTISQPRQVVLNFTDFQIGSSASCDTVIEI 876
Qy      568 RDGRTYAPLIGYHGTQVPOPLISTSNVLYLFTDKSHSDIGFQRLREYITLQSDHCL 627
Db      877 GPSSVLGSPNKEKFCSSNIPSFITSVNIILYVFVKSSMENHGFATKFSDXLE----- 931
Qy      628 DPGIPVNGQRHNDIFYGALVTPSCDGYTLSDSEPLECEBNPQMSALSCALCGGFI 687
Db      932 -----CGEVL 936
Qy      688 QGSAGTILSPGPDYFPNNLNCWTIETSHGKGVFFTEHTFLESCH---DYLLITENG 743
Db      937 TASTGIIESPGRVNVYRGVNCMTGCVVQGRQGLRLFEFSFYLEFGVNCNDYIEYDTA 996
Qy      744 SFTQPLRLQTLGSRILPADISAGLVNFTAVQVRFISDFSMYSYEGNITSEYDLEPCEBEV 803
Db      997 AQTF-LGRYCGKSIPELTS---NSNSIKLIFVSDSALAGEGFSINYEALD----- 1043
Qy      804 PAYSIRKGLQPGVGDITLFCFCFGRLEGTAARTICLGRRRIMSSPLRCVABEGNSVTG 863
Db      1044 -ASSV-----CLVYTD 1054
Qy      864 TQGTILSPNFPVNNNHECIYSIOTPGKGIQKARAFELSE-----GDVLKVYDGN 917
Db      1055 NFGTLSSPNFPNNYPSWBEICRTTVGLNQIALGFDTLLEDVFGSQCDVPVIRGCV 1114
Qy      918 NSARLGVFSHSEMGVTLNSTSSLWLDPTDAENTSKGPELHF----- 962
Db      1115 ETSPLVIGYGS-VLPTTIISGNKMLKFKSDALATAKGFSAVWDSSGTCGGNLTPPQ 1173
Qy      963 -----SSEFLICEDGTPKFGYKXNDHEHFAFGSSVSFSCDGY----- 1001
Db      1174 VLTSPNYPMPYSGSECYWRLSEAGSPPELEFODFLBEGP---SCSLDYGRVDGPT 1229
Qy      1002 -----SLRGESELLCLSGE-----RRTWDRPLPTVAEC 1030
Db      1230 TNSRLIDKLGGDTTPAIRISNKQVVLKTBELMGQGRGEINFRRCNVV----- 1282
Qy      1031 GGTIVGVSQVLSPGYPAPYEHNLCIWTIEAGCTTGLHFLVDPTEEV---HDVLR 1086
Db      1283 ---IVNKTFGILSINYPNPDYDKQRCNMWTIQATTGNTVYTFGLGFVESYMCSTDYVE 1339
Qy      1087 IMGPVVSQVLKELSPALPKDLHSTFNSVVLQFSTDFTS--KQGFALQFSVSTATSCN 1145
Db      1340 LYDPPQWNG---RYCGNNMPPGATTGSQLGVFGDGINSGKGRKO----- 1385
Qy      1146 DPGIPQNGSRSGDSEAGDSTVFOCDPGVALQSGAEISCVKIEHRFPWQSPPTCIAPCG 1205
Db      1386 -----WFTYCG-----CG 1392
Qy      1206 GDLTPSGVILSPNYPPEPPGKCDWKVTVSPDYIALVFNILFNLPG---YDFLHIY 1261
Db      1393 GEMGTAGSFSSPCYPSYFGNKECIWNIRVAPSSIQLTIGDPEVEGTSNCVDSLEIY 1452
Qy      1262 DGRPSLPLGSRFG---SOLPGRIBESSNSLPLAFSDASVSNAAGVIVIDYENRRSCF 1318
Db      1453 AGIDFNSPRIAQLCSGSPSANPMQVSTGDELARFTDSTLNGRGFNASRAVP----- 1507
Qy      1319 DPGSIKNGTRVGSDLKLSSSVTYCHGGEVEGHTSLSCILGPGKPVMMNPPPVCTAPC 1378
Db      1508 -----CG-----C 1510
Qy      1379 GGQYVSGDGVVLSPNYPONTSGQICLYPYTVPK-DYVFGQAFPHFALNDVVEVHDG 1437
Db      1511 GGIQLSRGIBSPNYPNNRYANTECSMIQIYVERGSRVLLNTIDFLEADPSCLRLMDGS 1570

```

QY 1438 SOHRLSLSGSHGTEGSLPLATSNOLIKFSAGKLAPARGEHFVQAVPTSATQSSV 1497
 Db 1571 SSTAPAAVAVCGROQPPNSITIASGNSLFRFRSSSSGNNRGRFAEFBE----- 1618
 QY 1498 PEPRYKRLGSDFSVGAIVREPCNSGYALQSPBEICLPVGAALQMNVAAPTCVPCGG 1557
 Db 1619 -----ECGG 1622
 QY 1558 N-LTERGTLSPGPPPYNSLNCWKI-VPEGAGIQIOVVSF-----VTEQWDSLEV 1611
 Db 1623 RIMTSDSDTLFSPLYPGNYLGNQCSWITEAOPFPNGITLSTFTGFOQNSTDCTRDVEI 1682
 QY 1612 FDGADNTVTLGSPGTTVPALLNSTNOULAFYSDISVSAAGFHELYTGLSCEP 1671
 Db 1683 LDGNDYAPVQGRGCGFSLPGLIISFGNALTIVRVTSTRSFEFRATYSA----- 1733
 QY 1672 AVPSNGVYKTERYLVNDVVSFOCEPBGALQHAHISCMPTVRMNPPLICACGGTV 1731
 Db 1734 -----STSSCGSF 1742
 QY 1732 EMEGVLLSPGPGNYPNNDCKIALPVGFAHIOFLNSTER-----NHDYETRNGP 1787
 Db 1743 YTLGRTNSPDYPADYGPNAECVNNIASPGRNLQSLFSLFNLSLNCNKDFVEIREG- 1801
 QY 1788 YETSRMNGRSGSELPSLLSTSHETT-VYFSDHSQNRPGFKLEYO----- 1833
 Db 1802 NATGGLIGRCQGNLPLNYSABEGSLMVRFVSDSGSTGMGFQARFNIRGNNNYVGTG 1861
 QY 1834 AYELOECPDEPFF-----ANGIVGAGYVNGQSVTFECL-----PGYOLTGH 1875
 Db 1862 KIASPFWPGKYFYNSNYKVVNNDAYGIIIGRILEMDIPEPTNCFYSLKTYGDFD-TGS 1920
 QY 1876 PVLTCQHT-----NRNW-----DHLPKCEVP---CGG 1901
 Db 1921 RLICTYGTGTESSTSSRNVLITFQSSDSVSGRGLIFEAADVSDSTPTTAPACGG 1980
 QY 1902 NITSSNGTV---YSPGPPSYSSQDCVWLITVPIGHVRLNLSLQTER---SCDFITI 1955
 Db 1981 FMVGTDPVGIFFSGMPREYANGADCIWIIYAP-DSTVELNITSLDIEPQGSNVDKLYI 2039
 QY 1956 WDBPQOAPRLGVTRMAKKTVOSSNOVLKFNHRAAGLPAFAFSAVPLTKCP--- 2012
 Db 2040 KDGSDLSPELAVLCGVSPGPFRSTGEYWIIRFTSSTVAG--TGFMASFGKSCGYI 2096
 QY 2013 -----PPTILPNAEVTENEENFICDIVRCLPFTL-----VGNELI 2051
 Db 2097 GADRGVITSPKYDVTYPLNLC-----SMGVLVQ---TGLTIAVGFQFPQIONRBS 2145
 QY 2052 TCKLGYLOF-----EGPPI-----CEVHCPT-----NELI----- 2078
 Db 2146 FCSQGDYLVLRNPGDNGSPILGSGRNGRFGCMYASTLTFTSGNEMFVQFISSNSGCG 2205
 QY 2079 -----TDSNGVILSOSYPGSPQFOTCSMLRVREPDVNLISLTVY 2118
 Db 2206 FKTRVKAISLACGTYITIGDADSDGYLTSBNYPAHYPQACBECIMLEAPGRGRIQOFE- 2264
 QY 2119 FLSEKDYDEFEIIFDGSGGSLPLKALSG-NYSAPLIV-----TSSSNSVYLKWS 2166
 Db 2265 -----DQENIEDTPEKSVSYELRDNANSNARLVSKLCCGTLPSGWSRRIRYLKNG 2317
 QY 2167 SDHAYNRKGFKIRYSAPYGLPAPLHGFILGOTSTOPGSGIHFGCNAGRYLVGSHMAIC 2226
 Db 2318 TDGSSSYMGFKAYSIASC-----GGTVGSDSGVLE----- 2348
 QY 2227 TRHPQVHLWSEALPLCOALSCGLPEAPKXMGVGEKYVTGTAAVY--SCSEGYHLOAGA 2284
 Db 2349 -----SIGYPTLPYANNVFCQFTRGLPGYLTLSFEDFNLOQSP 2388
 QY 2285 EATAECLDTGLWSN-----RNVPQCVPTCPDVSJ-----SVEHGRWLLIF 2327
 Db 2389 GCTKDFEVE--IWENGTSGRYLGRYCGNSTPSSVD--TSSNVASVXFVTDSDSVTASGRLEF 2445
 QY 2328 ETGYOFOAQLMLLCDGYYVYTGQRYVRCQANGKMSLGDSTPFCRIITISGELPIPPNGHRI 2387

Db 2446 KSRQ-----VCG-----GDIIGFTGT---FTSNYPNPNPGARI 2477
 QY 2388 --GLSVYATAIF-----SCNSGYTLVGSRYRECMANGLMSGSEVRCLAGHC 2433
 Db 2478 CEWTITVQGRRLVLFNTLRLSTOPSQNSGLIV-----FNQIRSNS----- 2520
 QY 2434 GTPEPIVNGHNGENSVYSGSVVYOCNAGFRLIGMSVRIQQ-----DHHMSKTPFCVP 2488
 Db 2521 ----PILQ-----KLSRYNVTNERKSGNTWKVVF 2547
 QY 2489 ITGHPGNVPVNGLTQGNQFNLDNVYFVCGNPGYMAEGARSCGLASGQMSDMPTGRIIN 2548
 Db 2548 FIDG--SRPYGFT-----ASTYEDAVCGG-----FLPSVGGN 2581
 QY 2549 CTDPGHQENSVR-----QVHAGCPHRSFGTTVSVRCHNGFYLLGTPLVSCQ----- 2595
 Db 2582 FSPGY--NGIRDYARNLDCEWTLISNPNRENSISIT-----YFLESTISGQDCTPD 2631
 QY 2596 -----GDGTMDRPRPQCLIVSCGHPGSPHQ-----MSGD--SYTVAAVRYG---CI 2639
 Db 2632 VLEFRVGDADGPLIEKFCSLAPAPLVIPYQVMIRFVSNRVEYT--GFYIEYSFTDG 2690
 QY 2640 GKRT---LYGNSTRMCGLDGHWTSGLPHCS-----GTSVGCDDPGI PAHGIRLGRS 2688
 Db 2691 GIRTGNGVSISSPNVNLYSAMTG-----CSWLKAPEGT-----ITL--- 2729
 QY 2689 FDPGTVMRPSGAGHYLRGSSERTCOANGMSGOSPECCVISCGNP-----GTPSNARVY 2743
 Db 2730 ----TLSDFLLEAGPTC--TSDSVTVRNGDSPGS--FVIGRY--CGQSVPRPIQSGSNQILIV 2781
 QY 2744 FSDGLVFESSIYEBREGYAT-----GLSRHCSVNGT-----WTGSDPE---CLVI 2788
 Db 2782 -----TFNTNNOGQTR--GFYATWTNNAJCCGTFPSANGTITSPCMPTFPNSRCSMTV 2835
 QY 2789 NCGDGPAPANGRLANDRRYN-KTVTYQCVPGYMMESHRSVLSCTKRTMNGTRPVCKA 2847
 Db 2836 ITGD-----SKWEISIFDSNFRIPSSDQONSFV-----KVMGRLMINKT 2877
 QY 2848 LMC-----KPPPLPNGKVVOSDPMGMSVYIACLEGYOLSPAVFTCE--GNGSWTG 2898
 Db 2878 LIAISCGDVAAPPIYTSNGI-----FTAVFQSEBMAAGFSAS 2915
 QY 2899 ELPOCFPVYFCGDPG---VPSRGRREDRGS-----YXSVSFSCHPPLVLYG 2942
 Db 2916 FISRGRTRFNSPGIISPNPKQYDNMNMCTYILDADPOSLVITFVPSFGLBDSALITG 2975
 QY 2943 SPRRFCQSDG-----TWSGTQ---PSCID-PTLTT-CADPGVPOFGIONN 2982
 Db 2976 T-----CDGDLGIIKGRMLNSTPLVTITGSETLRPLTVDGVLVLFYSDAYTTDFGFKIS 3031
 QY 2983 SGQYGVGSVTLFRCCQGYLLGOSTTRTCLPN 3013
 Db 3032 YRAITCGG--LYNESSGILRSPSYSYSNYPN 3060

RESULT 5
 US-09-612-314A-52
 ; Sequence 52, Application US/09612314A
 ; Patent No. 6713606
 ; GENERAL INFORMATION:
 ; APPLICANT: SMITH, RICHARD ANTHONY GODWIN
 ; APPLICANT: DODD, IAN
 ; APPLICANT: MOSSAKOWSKA, DANUTA EWA IRENA
 ; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
 ; MEMBRANE-BINDING AGENTS
 ; FILE REFERENCE: 37945-0004
 ; CURRENT APPLICATION NUMBER: US/09/612,314A
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: US 09/214,913
 ; PRIOR FILING DATE: 1999-03-16
 ; PRIOR APPLICATION NUMBER: PCT/EP97/03715
 ; PRIOR FILING DATE: 1997-07-08

PRIOR APPLICATION NUMBER: GB 96 148 71.3
PRIOR FILING DATE: 1996-07-15
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 52
LENGTH: 1947
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: CRI
US-09-612-314A-52

Query Match 8.1%; Score 1375; DB 4; Length 1947;
Best Local Similarity 22.0%; Pred. No. 3.5e-97;
Matches 540; Conservative 267; Mismatches 842; Indels 810; Gaps 109;

798 CEEBEVPA--YSIRKGLQFGVSDTLTFSCEPGYRLBGTARITC-LGRRRLMSSPLPRC 853
63 CRNPDPDVNGMVAHIKGIQF--GSQIKYSCTKGYRLIGSSSATCIIISGDTVIMNETPIC 120
854 -VAECGSVTCTGTLSPNPVNNNHCEIYSIQOPGKGIOLKARAFELSGDVLK 912
121 DRIPCGLPPTTNDGFISTN--RENFHYGSVVTYRCNPGSG--GKRVFEL-VGER-SI 172
913 YDGNNGARLLGVFSEMMGVTLNSTSSLMIDFIDAENTSXGFEHFSSFELICED 972
173 YCTSNDOQ--VGISGAPOCIIPN-----KCTP 199
973 PGTPEKFGYKVDGHP--AGSSVSPSCDPGYSLGSEBELLCISGERRTWDRPLPTVAEC 1030
200 PNVEN-GILVSDNLSFLNEVEFRCPGPFVKGPRVKQALNK--MEPELPSGRVC 256
1031 GGTGREGVSGVLSPGYAPAEHNLNCTIWEAAGCTIGLHFLVPTBEVHDVLRITWDG 1090
257 -----QPP-----DVLH----- 264
1091 PVESGVLLEKSLGAPLPDLHSTFNSVVLQFSTDFTSKQGFALQFSVSTATSCNDGIP 1150
265 -----A 265
1151 QNGSRGSDSWAGDSTVFOCDGYALOGSABISCVKIENRFWPSPTCTIAPCGDLTG 1210
266 ERTQRDKNFSPGQEVFVSCBPGYDLRGAASMRCTPOD--WSPAPTCGVKSCDDPMG 322
1211 P--SGVLSPPYPPFPKCEKDMKVSPDYVALVFNLENPEPGVDFLHIYGRSLS 1268
323 QLNLGRVLP-----VNLQLGAKVDFVC----- 346
1269 PLIGFYSGLPGRIESSNSLFLA-----FRSDASVNAGFVIDYENPRESCFDPGSI 1323
347 -----EGFOLKG--SSASVCYLAQMSLSNVSVPCEQIF-----CPSPFVI 386
1324 KNSTRVSSDLKL--GSSVTYYCH-----GGYVEGTSTLSCLIGDGKPVNNPPVC 1374
387 PNGHTGKPLEVFPFGKAVVNTCDPHDRGTSFDLIGESTIRCTSDQNGNWSPPARC 446
1375 TAPCGGQVSGDVVLSPNYFQNTSGQICLYFTVFKDYVVFQGFAPFTALNDVVEH 1434
447 -----GIL----- 449
1435 DGHSG--HSRLLSLSGSHGESLPLATSNQVLKFSKGLAPARGF-----HFVQAV 1486
450 -GHQADHFLFALTKQTMAADPFIQTS-----LKTGCRPEYGRPSITCLDLWSS- 503
1487 PR--TSAQCSSEVEPRYK-RIGSDPSVGAIVFECSNGVALOGSBEIECLPVGALAQ 1543
504 PKDVCXKRSCTPRDPVNGMVAHITDIQVGRINVSCTTGHRLLGHSSABET-LSGMAH 562
1544 MNVAPPTCV-VPCGSLTERGTILSPQEPPEYIANSLNCWKIYVPEGAGIQIOVSPVT 1602
563 WSTRPPTICQIPCGLPPTIANGDPISTNRENFHYS-----VVTYRC 604
1603 EQNWDSELEVFDGADNTVTLGSPSGTIVPALNSTSNQLYHFPYSDISVSAAGPHLEBKT 1662

605 NPGSGKRVF-----LVGDS-----ICTSND-----DQ 630
1663 VGLSSCEP--AVPS-----NGVTKGR--YLVNDVVSQCEGVALQHAHISC 1709
631 VGIMSGAPOCIIPNKCTPPNVENGILVSDNRSLSFLNEVEFRCPGPFVKGPRVKQ 690
1710 PGYVRBNVPPPLIAQCGGTVEEMEGVILSPGPNYPSNMOCMKIALPVGCAHIQF 1769
691 --ALNKPEELPSGRVC----- 706
1770 LNFSTBNHDYIIRNGPYETSRMMGRFSGSELPSLSLSTSHETTVFHSDHQNRBGF 1829
707 -----QPPPVVL-----HARTGR----- 720
1830 LEYQAYELQCPDPEPANGIVKAGYNVQGSVTFECLPGYQLTGHPVLTQCHTNAMD 1889
721 -----DKDN-----SPQGEVFSCEBPGYDLRGAASMRCT--PGDWS 756
1890 HPLPKCEVPCGNTSS--NGTVYSPGFPPYSSQDCWLLIYPIGHGVRNLNLSLQTE 1947
757 PAAPTCVKSCDDPMQGLNGRVLP-----VNLQLGAKV----- 791
1948 PSQDFTIWDGPOOTAPRLGVFTSRMAKTVOSSNQVLKFKHRDAATGIFALAFSAYP 2007
792 --DFVCD-BGFOLKG-----SSASCVL-----AGMESLMSNVSVP 823
2008 LTK--CPTTILPNAEVTEN--EENIGDIYRYCLP-----GPTLVGNEILLTKL-- 2055
824 VCEQIFCSPSPVPIPNGRHTKPLEVFPFGKAVVNTCDPHDRGTSFDLIGESTIRCTSDP 883
2056 --GTIYQFEGPPICEV--HCPNELLTSTGVILSISQSPGSPQOTGSMVLRVBPDY 2110
884 QNGVW--SSPARCGILGCOA-----PD- 906
2111 NISLTVYFLSEKQYDFEIFDGPSCQSLKALSGNYSAPLIVTSSNSVYLRMS--DH 2169
907 -----HFLFALKQTQTNASDPIGTSKLYEGRPEYGRFATICLDNLV--WSPKPD 956
2170 AYNRKGFIRYSAFYCSLPAPLHGLTGTSTQPGSHPFCGNAGYRLVGHSMATGRH 2229
957 VCKRKS-----CTPPDPVNGMVAHITDIQVGRINVSCTTGHRLLGHSSABECLS 1007
2230 PGYHLMSEALPICOALSGLPEAPKGMVFG--KEYVTGTAUVYSCSGYH--LQA 2282
1008 GNTAH-WSTRPPTICQIPCGLPPTIANGDPISTNRENFHYSVVTYRCNLGSRGVFEL 1066
2283 GAETAECL-----DTGLMS-----NENVP-----QCV 2306
1067 VGEBSIYCTSNDOQVIGMSGAPOCIIPNKCTPPNVENGILVSDNRSLSFLNEVEFRQ 1126
2307 P-----VTC-----PDVSSIS-----VEHGRWLLFETOYFOALMLICD 2342
1127 PGFVWKPRHVKQALNKWPELPSCSRVCOPPELHGHGHTSHQDNFPGQEVFYSCE 1186
2343 PGYTYTGQRIYRQANGKSLGDSPTCTRIISC-----GELPIPNGRIGTSLV-YGATA 2397
1187 PGYDLRGAASLHCTPGDWS--DEAPRCVAKSCDDFLGOL--PHGVLPPLMLQLGAKV 1241
2398 IFSCNSGTYLVGSRVRECMANG--LWSGEVRCIACHCTPPEIYVGHINGE--NYSY 2451
1242 SFVCDDEFRKLGSSVSHICVLGMRSLMNSVPRVCEHIFCNPRAIILNGRHTGTPSGDIPY 1301
2452 RGSVYQCN-----AGFRLIGMSVRIQQDDH--WSGKTPTC-VPTTGH----- 2493
1302 GKEISYTCDBHPDGMFTNLIGESTIRCTSDPHNGVWSSPAPRCSELSVAAGHCKTPEQF 1361
2494 -----PGNPVUGLQGNQFNLDVVKYVCNGVMAEBAABOCLASGQMSMLPTCRINC 2549
1362 PPAFTPIINDP--EPVGTSLNECRGTYGKKFESIS-CLENLWSSVEDNCRKSC 1416
2550 TDGCHQENSROYVNASGPRHFSFGTIVSYCNHGFYLLGTPVLSQCGD--TWDRRPQ 2606

1417 GPPPEFNGVHINTD-----TQFSTVYVNSCNEBFRLLGSPSTCLVSGNNVTDKAP1 1472
QY 2607 CLVSCGHPSPSPHSQSGDST-----VGAVVRYSC-----IGKRLTVGN 2647
Db 1473 CEIISC-----BPPETISNGDFYSNNKRTSFHNGTVVYQCHTGPDGEQLFELVGRSIV-- 1526
QY 2648 STRMCGLDGHTGSLPHSCSTGVGCGDPGI PAHIGLGDG---FDPGTWRFSCGAGV 2704
Db 1527 CTSDDDGVWSSPPRCISTN--KCTAPFV--EVAIKVGNRSPFSLTEIRFCQGFV 1583
QY 2705 LRGSSERTQANGWSSSQPECGVISCNPGTSPMAVVS--DGLVFSSSIYVECHGY 2762
Db 1584 MVGSHTVQCOTNGRMGPELPHCSRV--CQPPFELHGHNTLSHQNFSPGQEVFVSCSPSY 1642
QY 2763 YATLLRHCSTVNTWTSDECLVINGD--PGIPANGLRGNDPFRKNTVYQCVGY 2820
Db 1643 DLRGASLHCTPQGDWSPPEARCTVKSODFLQGLPFGRYLLPLINTQLGAKVSFVCDGEGF 1702
QY 2821 MMESHRSV--VLSTCKRTWNGTKPVCKALMKCPPLIPNGKVVS--DEMMGSSVTV 2875
Db 1703 RLKGRSASHCLAGMK--ALMNSVVPVCEQIFCENPRALINGRHTGTFRGDIPIYKEISIA 1761
QY 2876 C-----LEGYLSLPAVFTCEGNSWTGELPQC--PVPFGDGVPSRGRRED 2921
Db 1762 CDTHPRGTMFNLIG--ESSIRCTSDPQNGWSSPABRCESVPAACFHPKXIONHYIG 1820
QY 2922 RGFSY---RSSVSPSCHPPLVVSPPRPFOSDGTWSTGTOPSCIDPLTLTTCADPGVPQ 2978
Db 1821 GHVSLVPGMTISTTCDPGYLVGKGFTFCTDQISWQLDHYCE--VNCSEFP--LPMNG 1876
QY 2979 IONN---SOGYOVSTVLFRCQKGYLLQGSTTRTCLFNLTWSTGTPDCV--PHHCRQPE 3032
Db 1877 ISKLEMKKYVHYGDYVTLKCEDGYTLGSGPWSGCCQADRMDDPLAKCTSRACICDGPX 1935

RESULT 6
5256642-2
PATENT NO. 5256642
APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
MINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F., JR., STEPHEN
H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
RECEPTOR 1 (CRL1) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
USE THEREOF
NUMBER OF SEQUENCES: 30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,128
FILING DATE: 24-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 412,745
FILING DATE: 26-SEP-1989
APPLICATION NUMBER: 332,865
FILING DATE: 03-APR-1989
APPLICATION NUMBER: 176,532
FILING DATE: 01-APR-1988
SEQ ID NO:2
LENGTH: 2039
5256642-2

Query Match 8.0%; Score 1362; DB 6; Length 2039;
Best Local Similarity 22.0%; Pled. No. 3,9e-96;
Matches 537; Conservative 266; Mismatches 833; Indels 810; Gaps 109;

QY 798 CEEPEVPA--YSIRKLOFGVGDTLTFCFPGYRLGRTAGTATTC--LGGRRRLMSSFLPRC 853
Db 104 CRNPPRPVGMVHVIQGF--GSQIKYECTKRYLIGSSSATCIISGTVIWMDETPIIC 161
QY 854 -VAECGNSVTGOTLLSPNFVYNNHNECIYISQTOFGKGIQLKARAFELSEGDVLKV 912
Db 162 DRIPCLPPLITINGDIFSTN--RENFHYGSVVTYACNPGSG--GRKVFEL--VGP--SI 213
QY 913 YDGNNSARLLGVFSHEMMGVTLNLTSSIMLDTTDAENTSKGELHFFSEFLIKED 972

214 YCTSNDDQ--VGWSGAPDQCIIPN-----KCTP 240
QY 973 PGTPEGYVYHDEGHP--AGSSVSFCDDPGYALRGSEELLCTSGERRTDRLPTCAVC 1030
Db 241 PVNEN--GLIVSDNRSLFSLNEVVERRCQVFPVFMKPRARKCALNK--WEPFLPSCSRVC 297
QY 1031 GGIVRGSEVGOVLSRPYPAPEHNLNCTWTIEAGCTTGLHPLVFDTEVHDVLRIMDG 1090
Db 298 -----QPPP-----DVLL----- 305
QY 1091 PVESVLLKELSGPALPKDLHSTFNSVVLQFSTDFTSKQFALQFVSSTATSCNDPGR 1150
Db 306 -----A 306
QY 1151 QNGSRGDSWEAGDSTVFOCDGYALQGSABEISCVKIENRFPWOPSPTCIAPCGDLTG 1210
Db 307 ERTQDKDNFSPQGEVYFSCBPGYDLRGASMRCTPQGD--WSPAAPFCEVKSODDFMG 363
QY 1211 P--SGVILSPNYPPEYPRPKCEDMKVTVSPDYIALVFNIFVLEPGYDFLHIYDGRDLS 1268
Db 364 QLLNRGRVLP-----VNLQUGAKVDFVCD----- 387
QY 1269 PLIGSFYSGQLPGRIESSNSLFLA-----FRSDASVSNAGFVIDYTENPRESCFDPGSI 1323
Db 388 -----EGFQLKG--SSASYCVLAGMESLWMSVVPVCEQIF-----CSPFVI 427
QY 1324 KNGTVGSDLKU--GSSVTVYCH-----GGYEVEGTSTLSCTLPGDKPVMNNPRVC 1374
Db 428 PGRHTGTPLFVFPFGKAVNYCDHPDRGTSFDLIGESTIRCTSDPQNGWSSPABRC 487
QY 1375 TAPCGQYVYSGDGVVLSNYPQNTYSGQICLYFVTVPPDYVFGQFAFHIALNDVBEH 1434
Db 488 -----GIL----- 490
QY 1435 DGHQ--HSRLSLSGHTESSPLATSNQVLIKFSAGLAPARGF-----HFVQAV 1486
Db 491 -GHQAPDHPFLFAKTKQTNASDPIGTS-----LKYECRPEYGRPFSTICLDNLVWSS- 544
QY 1487 PR--TSAQCSVPEPRYRGK--RLGSDSVGAIIVRECNISGALQSPETICELPFGALAQ 1543
Db 545 PPDVCKRKSCTPPDPVGMVHVTIDQVGRINVSCTTGRLIGHSSABECI--LSGNAH 603
QY 1544 MNVSAPTCV--VPCGNLTERGTILSPGFPBYLNSLNCWKIIVPEGAQIQVVSVTV 1602
Db 604 MSTKPIQORIPCGLPPTIANGDPISTNRENFHGS-----VITIRC 645
QY 1603 EONMDSLEVPDADNTVTMLGSFGTTPALLNSTSNQLYLHFYSIDISVSAAGFLBXYKT 1662
Db 646 NPGSGRKVFE-----LVGEPS-----LYCTSD-----DQ 671
QY 1663 VGLSCPEP--AVPS-----NGYKTGER--YLVNDVYSFOCEPGYALQGHANISCM 1709
Db 672 VGWISGAPQCIIPKCTPPNVENGILVSDRSLFSLNEVVERFCQDPFVWKGPRRVYCO 731
QY 1710 PGTVRNMYPPPLCIAQCGTVEEMEGVILSPGPGNYPNMDCSMKIALPVGCAHIQF 1769
Db 732 --ALNKWPELPSGSRV----- 747
QY 1770 LNFSTPEHNDYIETIRNGDEYTSRMNGRPSGSELSLSTSHETTVYFHSDSQNRDPFK 1829
Db 748 -----QPPDVL-----HAERTQ----- 761
QY 1830 LEQAYELQECRDPDEPPANGIVRAGVNGOSVTFECLPGQLTGHVLTQHOHTNNRD 1889
Db 762 -----DKDNF-----SPQGEVYSCBPGYDLRGASMRCT--POGDWS 797
QY 1890 HPLPKCEVPCGANTTS--NGTVYSPGFPSPBSSSODCWLITVPIGHVYLNLSLQTE 1947
Db 798 PAAPFCEVKSODDFMGQLNGRVLFP-----VNLQUGAKV----- 832
QY 1948 PSQGFITTWGPPQOTARLGLVFTRSMAKKTVOSSNOVLKFRDAATGCIPTAIFASAYP 2007
Db 833 --DFVCD--EGFQLKG-----SSASYCVL-----AGMESLWNSVP 864

QY 2008 LTK---CPPPTILPNAEVVTEN--EENIGDIIVRYRCLP-----GTLVNGEILLTCL-- 2055
 DB 865 VCEBIPFBSPBPVITNGHRTKPLVFPFGKAVANTTCDHPDRGRSFDLIGSTTRCISDP 924
 QY 2056 ---GTYLOFEGPBPICEV--HCPTNELTDTSTGVLSSQSYPSYPOFOTCSMLVVRPDY 2110
 DB 925 QGNGVW---SSPARRCGILGHCOA-----PD- 947
 QY 2111 NISLVEYFLSEKQYDEFEIFDGSQSPILKALSGNYSADLIYSSNSVYLWSS- DH 2169
 DB 948 -----HFLPAKLTQTNASDPFGTSLKYECPREYGGPFSITCLDNLV---WSSPKD 997
 QY 2170 AYNRKGRIRYSAYVCSLPAPLHGFILGOSTPGSINHGCAGYRLVHSHMALTRH 2229
 DB 998 VCKRKS-----CTPPDPVNGMVHVLIDIVGSRINSCITGHRLLIGHSSAECILS 1048
 QY 2230 POGYHLMSEALPLCOALSCGLPEAPKNGMVG--KEYTVGTAKVYSCSEGYH---LOA 2282
 DB 1049 GNTAH-WSTKPLICQRIPCGLPPTIANGDPISTRENFHOSVYTYRGNLGRKXYPFL 1107
 QY 2283 GAETABCL---DTGLMS-----NRNVP-----OCV 2306
 DB 1108 VGEPSIYCTSDNDQVIGWGPAPCIIIPNKCTPENVENGILVSDNRSLFSLNEVVDPRCQ 1167
 QY 2307 P-----VTC-----PDVSSIS-----VEHGRMLIFETQYQOQADMLICD 2342
 DB 1168 PGFVMPKPRRYKCALNKBEBELPSCSRVCOPPEILLHGEHPHONFSGQVFSYCE 1227
 QY 2343 PGYVYTGQAVIRCOANGKWSLGDSTPTCRILISC---GELPIPNHGRIGTASY-YGATA 2397
 DB 1228 PGYDLRGAASLHCTPGQDMS--PEAPRCAYVSCDPIFGLQ---PHGVLPPPLNQLGAKV 1282
 QY 2398 IFSCNSGYTLVGSFVRECMANG--LWSEVRCLAGHCTPEPIVNGHNGE---NYSY 2451
 DB 1283 SFVDEBFRLKGGSSVSHCVLWGMRSILWNNSVPCVCEHIFCPMPAILNGRHTGTPSGDIPY 1342
 QY 2452 RGSVYVYOCN-----AGFRLIGMSVRICQODH-----WSGKTIPC-VPIICG----- 2493
 DB 1343 GKELSTYCDHPDRGMTFNLIGESTIRCTSDPHONGVWSSPAPRCESLVRAGHCKTPEQF 1402
 QY 2494 ---PCNPVNGLTQGNQFNLDVYKFCNPGYMEGAARSOCLASGQMSDMLPCRIINC 2549
 DB 1403 PRASPTIPIINDF---EPVOTSLNYECRPRYBGMKMSIS-CLENLWSSVEDNCRKSC 1457
 QY 2550 TDRPHQNSVYRQVHASGPHRFSFGTYVSYRNMHGYLLGTPLVLSQCGDG---TWDRBPQ 2606
 DB 1458 GPPPEPFGMVHINTD---TQFSTVNVSCNEGFRILIGSPSTCLVSGNNVTWDKKAPI 1513
 QY 2607 CLVSCGHPPSPHISQMSGDSYT-----VGANVVRISC-----IGKRTLVGN 2647
 DB 1514 CEIISIC---EPPTJISNGDFYSNNRTSFMHNGTVVTVQCHTGPDEQLFELVGRSRTY-- 1567
 QY 2648 STRMCGLDGHWITGSLPHCSGTSVVCDDPGIPAHGIRLGDs---FDGTVMARFSCBAGHV 2704
 DB 1568 CTSGDDQVGVWSSPFRICISTN--KCTAPREV-ENAIIVPGKRSFSLTEIIRFCQGFV 1624
 QY 2705 LRGSSEFTCOANGSSWSQPECGVISCNPCTPSNARVFS--DGLVSSSIYVECEGY 2762
 DB 1625 MVSGHVTQCCQNGRMGPKLPHCSRV-COPPEILHGHHTLSHQNFSPGQEVFVSCBPSY 1683
 QY 2763 YATGLSLRHCSVNGTWGSDPECLVINGD--PGIPANGRLANDFYNKTVTYQCPVGY 2820
 DB 1684 DLRGAAHLCTPPQDMSPEARCTVKSCDPIFGLQPLPGRVLLPLNLDLGAKVSPVCEBGF 1743
 QY 2821 MIESHRVS--VLSCTKRTWNGTKPVCKALMKCPPLIPNGKVVS---DFMMSSTVYA 2875
 DB 1744 RLKGRASHCVLACMK-ALMNSVYVCEQIFCPNPRAILNGRHTGTFGDIPIYKELESTA 1802
 QY 2876 C-----LEGVLSLPAVFTCEGNSWTGELPQC---PFVFGDGPVPSRGRRD 2921
 DB 1803 CDTHPDRCMTFNLIG-ESSIRCTSDPQNGVWSSPAPRCESLVPAACPHPKIONGHYIG 1861

QY 2922 RGSFY---RSSVFSCHPLVLVGSPPRRFCQSDGTMSGTQPSCIDPTLTTCADPGVPOFG 2978
 DB 1862 GHVSLVLPKMTISTYTCDPGVLVWGKGFICTDQISMSQLHYCKE---VNCSEF-LPMNG 1917
 QY 2979 IONN---SOGYOVGSVTLFRCOKGYLLQGSTTRTCLPMLTWSGTP 3021
 DB 1918 ISKELMKVYHGDVYTLKCBQGYTLGSPWSCQCADRMD--P 1961
 RESULT 7
 5472939-2
 Patent No. 5472939
 APPLICANT: PEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
 MINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F., JR, STEPHEN
 H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.
 TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
 MEDIATED DISORDERS
 NUMBER OF SEQUENCES: 30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/138,825
 FILING DATE: 19-OCT-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 588,128
 FILING DATE: 24-SEP-1990
 APPLICATION NUMBER: 412,745
 FILING DATE: 26-SEP-1989
 APPLICATION NUMBER: 332,865
 FILING DATE: 03-APR-1989
 APPLICATION NUMBER: 176,532
 FILING DATE: 01-APR-1988
 SEQ ID NO:2
 LENGTH: 2039
 5472939-2
 Query Match 8.0%; Score 1362; DB 6; Length 2039;
 Best Local Similarity 22.0%; Pred. No. 3.9e-96;
 Matches 537; Conservative 266; Mismatches 833; Indels 810; Gaps 109;
 QY 798 CEEBEVDA---YIRKGLQFGVDTLTFCFPGYREGTARITC-LGRRRLMSSPLPRC 853
 DB 104 CRNPDPVNGMVHVIKIQF--GSOIKYCTKGYRLIGSSASATCIISGDVIMDNETPIC 161
 QY 854 -VAECNSVTCGTGLTSLPFPVNNNNHNCIYSIQOPKGIOLKARAFELSEGDVLK 912
 DB 162 DRIPCGLPITTTGDFLSTN---RENPHYSVVTYRCNPSG---GRKVEL-VGEP-SI 213
 QY 913 YDGNNSARLLGVFHSHEMGTLSSTSLMDFTIDAENTSKEFELHSSFELICED 972
 DB 214 YCTSNDDQ--VGIMSGAPRCIIPN-----KCTP 240
 QY 973 PGTPKGYKVHDEGHR--AGSSVSFSCDPGYSLRGSEELLCLSGERTWDRPLPTCYAEC 1030
 DB 241 PNVEN-GILVSDNRSLFSLNEVVEFRCPQVFMKGPBRVYKQALNK--MEPELPSCSRV 297
 QY 1031 GGVTRGVSQVLSPGVPARYENLNLCIWTIEAAGCTIGLHVLVPTDESVDHLYLRWD 1090
 DB 298 -----QPP-----DVLLH--- 305
 QY 1091 PVESGVLLKELSGPALPKDLHSTFNSVVLQFTDFTSKQGFALQFSVSTATSCNDPGIP 1150
 DB 306 -----A 306
 QY 1151 QNGSRSGDSWEAGDSTVFOCDPYALQSAEISCVKIENFFWQPSPTCIACGGDLTG 1210
 DB 307 ERTQORDKDNFSPQGEVYFSCPEGYDLRGAASMRCTPGD---WSPAPTCVKSCDPFMG 363
 QY 1211 P--SGVILSPNYPEPFPKGCDCMKTVASDYIALVFNFLENPEPGVDPLHITDGRDLS 1268
 DB 364 QLINGRVLP-----VNLQAKVDVCD----- 387
 QY 1269 PLIGFYSQLPRISSNSNLFIA---FRSDASVSNAGFIADYTENRRESCFPDGS 1323
 DB 388 -----EGPQLKG---SSASVYVLAQMESLMSNSVYVCEQIF-----CPSPPI 427

QY 1324 KNGTRVSGDJKL---GSSVTYCH-----GGYVEGTSTLSCILGPDGKPVNNPPVC 1374
 DB 428 PNRGHTCKPLEVFPFGAVNYTCDPHDRGTSPLISESTIRCTSDQNGVSSPAPRC 487
 QY 1375 TAPCGGYVSGDVLSPPNYPONTSGOICLYFVTVBKDYVVGQAFHTALNDVEVA 1434
 DB 488 -----GIL----- 490
 QY 1435 DGHSG---HSRLSSLSGSHGESLPLATSNQVLKFSAGKLAAPRGF-----HPVQAV 1486
 DB 491 -GHQADPHFLPALKTQTNASDPFGTS-----LKRECPREYGRPSITCLDNLWSS- 544
 QY 1487 PR---TSATQCSSVPEPRYK-RLGSDFSVGAIVAFECNSGVALQGSPEIECLPVPALAQ 1543
 DB 545 PKOVCKRKSCKTPRDPVNGWVHVITDIOVSRIYSCCTGHRLLGHSSACCT--LSGNAH 603
 QY 1544 MNVSAPTCV-VPCCGNLTERRGITLSPGPEPYLNSLNCWKIVPEGAGIQIVVSFT 1602
 DB 604 MSTPPIQRIPCGLPPTIANGDFISTNRENFHGS-----VVTYRC 645
 QY 1603 EQMWDSELEVDGADNTYMLGSPGTTVPALLNSTQVLYHFIYSIDISVSAAGFHEXT 1662
 DB 646 NPSGGKRYPE-----LVGEBP-----LYCTSD-----DQ 671
 QY 1663 VGLSSCEP- -AVPS-----NGVKTGER- -YLVNDVVSFOCEBVALQCHAHISCM 1709
 DB 672 VGIWSGAPFOCIIPNKCTPPNVENGILVSDNRSLFSLNEVFERCQGFPMKPRVKQ 731
 QY 1710 PGTVRKNVPPPLCIAQCGGTVEBMEGVILSPGPGNYPSNMDCKIALPVFGAHIQP 1769
 DB 732 - -ALNKPEPLPSCSRVC----- 747
 QY 1770 LNFSTERNHYIELRNQPYETSMNGRFSGSELBSLSLSTSHETTYFHSQNRGPK 1829
 DB 748 -----QPPDVL-----HAERTQ----- 761
 QY 1830 LEQVAYELQECRDEPFPANGIVGAGYVQSVTFECLPGYQUTGHPVLTCQHATNRND 1889
 DB 762 -----DKDN-----SPQEVFYSCEPQYDLRGAASMCCT--PGDMS 797
 QY 1890 HPLPKCEVPCGAGNITS- -NGTVYSPGFPSPYSSQDCWLITVPILGHGVRLNLSLQTE 1947
 DB 798 PAAPTECVKSCDPMQGLNGRVLFP-----VNLQGAKY----- 832
 QY 1948 PSGDFTIWDGPOQTARLVFTFRMAKKTVOSSSNQVLKFRHDATGSIFAIASAVP 2007
 DB 833 ---DFVCD-BGFQUG-----SSASTCVL-----AGMESLWSSVP 864
 QY 2008 LTR---CPPTILPNAEVVTEN-EEFNIGDIVRYCLP-----GFTLVGNEILTCKL-- 2055
 DB 865 VCSQIFCPSPVPIPNRHTGKPLEVFPFGAVNYTCDPHDRGTSPLIGESTIRCTSDP 924
 QY 2056 ---GTIYQFEGRPICEV- -HCFPTNELDSTGVILSOSYFGSPQFQTCMWLVREPDY 2110
 DB 925 QGNGVW---SSPAPRCGILGHQA-----PD- 947
 QY 2111 NISLVEYFSEKQYDFEILFDGSPGOSPLKALSGNYSAPLIVTSSNSVYLWMS-DH 2169
 DB 948 -----HFLFALKTQTNASDPFGITSLKYBCRPYEGFPSTICLDNLV--WSPKPD 997
 QY 2170 AVNRKGFRIYASAPYSLPRAPLHGLIGTOSTOPGSLHFGNAGRYLVGSHMACTH 2229
 DB 998 VCKRKS-----CTRPDPVNGWVHVITDIOVSRIYSCCTGHRLLGHSSACIIS 1048
 QY 2230 POGYHLMSEAIPLCOALSCGLPEAPKNGVFG--KEYTVGYAVYSCSEGYH---LQA 2282
 DB 1049 GNTAH-MSTKPIQRIPCGLPPTIANGDFISTNRENFHGSVVTYRCNLGSGRKYFEL 1107
 QY 2283 GABATAECL---DTGLMS-----NNVNP-----QCV 2306
 DB 1108 VGEPSYICTSNDQVGIWSGAPFOCIIPNKCTPPNVENGILVSDNRSLFSLNEVVDPRQ 1167

QY 2307 P-----VTC-----PDVSSIS-----VEHGRMLIFETOYQFOALMLICD 2342
 DB 1168 PGFWMKPRARVACQALNKKPELPPSCSRVCQPPPELHBEHTRPSQDEYVYSSE 1227
 QY 2343 PGYYTGORVIRCOANGKWSLSDSTPTCRIISC---GELPIEPNGHRIGLTSLV-YGATA 2397
 DB 1228 PGYDLRGASLHCTPGDWS--PEARCAVKSCDPLGOL---PHGRVFLPMLQGAKY 1282
 QY 2398 IFSNCGTYLVSRRRECMANG---LMSGSEVRLAGHCTGTEPIYNGHINE--NYG 2451
 DB 1283 SFVCEGFRKGSVSHCVLVMRSLWNNSVPEVCEHIFCPNPALINGHTGPSGDIY 1342
 QY 2452 RGSVYQCN-----AGFRLIGMSVRI-COODH-----MSGKTPFC-VPIYTCG----- 2493
 DB 1343 GKEISTCDPHDRGMTFPLIGESTIRCTSDPHGNGVSSPAPRCESLVRASHCKTPEQF 1402
 QY 2494 ---PENPVNGLTQGNQFNLDNVKFCVCPGMAEGARSOCLASQWSDMLPTCRIINC 2549
 DB 1403 PRASPPIPNDF-----EPFVGTSLNYECRPGYFGKMFIS- -CLEMLWMSVEDNCRKSC 1457
 QY 2550 TDPGHENSVRQVHASGHRFSGTTVSRCNHGFVLCTPVLSCQDGC---TWDRPQ 2606
 DB 1458 GPPEPEFNGMHIND- -TQGSTVNYSCNEGFRLLGSPSTCLVSGNNVTDKKA 1513
 QY 2607 CLVSCGHPSPPHSQMSGDSYT-----VGAVVRYSC-----IGKRTLVGN 2647
 DB 1514 CEIISC---EPPTISNDPFSNNRTSFHNGTIVTYOCHTPDGEQLFELVGERIY-- 1567
 QY 2648 STRMCGLDGHTGSLPHCSGISVGVCGDPIPAHIGRLGDS--FDPGTVMRPSCAGHV 2704
 DB 1568 CTSKDQVGVWSSPPRCISTN- -KTAPEV-ENAIRVGNKSPFSLTEIIRFCQPGFV 1624
 QY 2705 LRGSEHRTQOANGSNGSQPEGCVISCGNPGTRSNARVVS--DGLVFSSTIYVEGREY 2762
 DB 1625 MWGSHTVQQTNGRNGPKLPHCSRV- -COPPELHGEHTLSHQDNFSPQOEYFISCEPBY 1683
 QY 2763 VATGLLSRCHSVNGTWSDEPCLVINGD--PGIPANGRLGDNFRVKNKTYTQCVPY 2820
 DB 1684 DLRGASLHCTPQGDWSPEARCTVKSODPLGQPLHGCVLLPMLQGAKYSPVCDDEG 1743
 QY 2821 MMESHRSV--VLSCTKDRTWNGTKVYCALMKKPPPLIPNGKVVS--DFMMSGVYA 2875
 DB 1744 RLKGRSASHCVLAGMK-ALMNSVPEVCEQIFCPNPALINGHHTGPPGDIYGEKISYA 1802
 QY 2876 C-----LEQVYSLPAVFTCEGNGSWTGLPQC---FPVFCGPGVPSRGRARD 2921
 DB 1803 CTHPDRCMTNPLIG-ESSIRCTSDPQNGVSSPAPRCESLVPAACPHPKIQNGHYIG 1861
 QY 2922 RGSFY---RSSYFSCHPPLVYVSPRRFCOSDGTWSTGTOPSCIDPTLTTCADPGVPOG 2978
 DB 1862 GHVSLYLPQMTIYSTCDPEYLLVGKFICTDQGLMSQLDHYCKE---VNGSFP-LFMNG 1917
 QY 2979 IONN---SOGYQVGSYVLFRCQKGYLLOGSTTRTCLPMLTWGTPP 3021
 DB 1918 ISKELEMKKVHYHGYVTLKCEDGYTLRGSFWSQOADDNRD--BP 1961

RESULT 8
 5256642-10
 Patent No. 5256642
 APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
 WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
 H.; MAKRIDES, SAVVAS; MARSH, HENRY C.; JR.
 TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
 RECEPTOR 1 (C1R1) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
 USE THEREOF
 NUMBER OF SEQUENCES: 30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/568,128
 FILING DATE: 24-SEP-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 412,745
 FILING DATE: 26-SEP-1989

APPLICATION NUMBER: 332, 865
 FILING DATE: 03-APR-1989
 APPLICATION NUMBER: 176, 532
 FILING DATE: 01-APR-1988
 SEQ ID NO: 10
 LENGTH: 1847
 5256642-10

Query Match 7.5%; Score 1267.5; DB 6; Length 1847;
 Best Local Similarity 21.7%; Pred. No. 7.6e-89;
 Matches 503; Conservative 253; Mismatches 768; Indels 795; Gaps 103;

798 CEEBEVA---YSIRKGLQFGVDTLTFSCPPGYRLGSTARITC-LGRRRLMSSPLPRC 853
 109 CRNPDPVNGMVHVIKGIQF--GSQIKSCCKGYRLIGSSATIGSDIYIMNENRPIIC 166
 854 -VAECGNSVTGCTGLISPNFVNUNNNHCTIYSIQTPGKIQLKARAFELSGDVLC 912
 167 DRIPCGLPRTITNDPFISTN---RENPHYGSVVTYRCNPGSG--GRKVFEL-VGEP-SI 218
 913 YDGNNSARLLGVSHBMMGVTLNSTSSLMDFITDAENTSKGFELHFSFELICED 972
 219 YCTSNDOQ--VGISGAPPOCTIPN-----KCTP 245
 973 PGTPEFYKYHDEGHF--AGSSVSFSCDPGYSLRSEELCLSGERRTWDRPLPTVAEC 1030
 246 PNVEN-GILVSDNLSFLSNEVVEFRQCPVPMKGPRAVKQALNK--WEPELPSGRVC 302
 1031 GGTVRGVSQVLSPGYPAPYEHNLNCITWIEAGCTIGHLVLPTEVHYDLRIWDG 1090
 303 -----OPP-----DVH----- 310
 1091 PVESGVLLKELSGPALPKDLHSTFNSVLFQSTDFTSKQGFALQFSVSTATSCNDGP 1150
 311 -----A 311
 1151 QNGSRGDSWEAGDSTVFOCDPGYALQSAEISCVKIEENRFMQSPPTCIAPCGDLTG 1210
 312 ERTRQDNDNPSPOGEVFPYSCPGYDLRGAASMRCTPGD--WSPAPTCGVKSCDDPMG 368
 1211 P--SGVILSPNYPEPPYPPGKECDMKVTSPDYIALVFNIFNLEPGYDFLHIYDGRDLS 1268
 369 QLNGRVLPF-----VNQLGAKVDFVCD----- 392
 1269 PLIGSFYGSOLPGRIESSNSLFLA-----PRDSASVSMAGFVDYDYNREPSFDPGSI 1333
 393 -----EGFOLKG--SSASVCVLAGESSLWNSSVPPCEQIF-----CPSPVI 432
 1324 KNGTRVGSDELK---GSSVTYYCH-----GYVEGSTLTCILGDPKPVNNPPRVC 1374
 433 PNGRHTKPLLEVFPFGKAVNYTCDPHDRGTSFPLIGESTIRCTSDPOGNGWSSPARC 492
 1375 TAPCGGYGSDGVLSPNYPONTYSGOICLYFTVPEKDYVVFQAFPHITALDVEVH 1434
 493 -----GIL----- 495
 1435 DGHSG--HSRLLSSLSGSHGSELPLATSNQVILKFAKGLAPARGF-----HFVQAV 1486
 496 -GHQADHFLPALKQTUNASDPFIST-----LKYECREPYRPRPSITCLDLWSS- 549
 1487 PR--TSATQCSSVDEPRYK--RLGSDFSVGAIVFECNSGYALQSGPEIECLPVGALAQ 1543
 550 PKOVCKRKSCKTPDPVNGMVHVIKGIQF--GSQIKSCCKGYRLIGSSATIGSDIYIMNENRPIIC 166
 1544 MNVASPTCV--VPCGNLTERRGTLSPGFEPYVINSINCWKIVPBGAGIQIQVASFV 1602
 609 WSTRPPIQRIPIPCLPPIANGDFISTRENPHYS-----VVTYRC 650
 1603 EQMNDLSLEVPGANNTYMLGSPFGTIVPALNNTSQVLYHFYSDSISVSAAGHLEYKT 1662
 651 NPGSGKRVFE-----LVGEPS-----TYCTSD-----DQ 676
 1663 VGLSSCPPE--AVPS-----NGVKTGER--YLVNDVVSFQCEPGYALQGHANISCM 1709

677 VGIWSGAPPOCTIPNKTTPNVEKGLVSDNRSLSFLNEVVEFRQCPGVKGRARYKQ 736
 1710 PGTVRNNYPPPLDIAOCGGTVEMEGVILSPGFPYPSNMDCSKIALPVGGAHIQF 1769
 737 --ALNKWEPELPSGRVC----- 752
 1770 LNFSTENHDIYELRNQYFENSRMGRFSGSELPSLSLSTSHETTYFHSQSNRPGFK 1829
 753 -----QPPDVL-----HETTR----- 766
 1830 LEYQAVELQRCPPDEPRANGIVRGAGVNSQVTFECLPGYOLTGHPVLTQCHQTNRMND 1889
 767 -----DKDNF-----SPQGEVYSCBPGYDLRGAASMRCT--PQDWS 802
 1890 HPLPKCEVPCGNGITSS--NGTYVSPGFPPSYSSQDCWMLITVPIGHVRLNLSLQTE 1947
 803 PAAPTCGVKSCDDPFMGQLNGRVLPF-----VNQLGAKV----- 837
 1948 PSGDFITIMDPOQTARLGVFTSMMAKTVQSSSNQVLLKFHRDAATGGIFATAPGAP 2007
 838 ---DFVCD-BGFOLKG-----SSASTCVL-----AGNESLMSNV 869
 2008 LTK---CPPTILFNAEVTEN--EEFNIGDIVRYRCLP-----GFTLVGNEILTCRL-- 2055
 870 VCBQIFCPSPFVIPNGRHTKPLLEVFPFGKAVNYTCDPHDRGTSFPLIGESTIRCTSDP 929
 2056 ---GTYLQFBGPPPIGV--HCPTELLTDSGTVILSOSYPSGYPOFQTSWLVVERPDY 2110
 930 QNGVW---SSPARCGLIGHCOA-----PD- 952
 2111 NISLTVYFLSEKQYDFELFDGSGOSPLLKALSGVNSAPLYVTSSNSGYLWSS--DH 2169
 953 -----HFLPAKKTQTNASDPFISTLYECREPYRPRPSITCLDNLV--WSPKX 1002
 2170 AYNRKFKIRYSAPYCSLPRAPLHGFILQGTSTPGGSIHFGNAGRYLVGSHMAISTRH 2229
 1003 VCKRKS-----CTRPDPVNGMVHVIKGIQF--GSQIKSCCKGYRLIGSSATIGSDIYIMNENRPIIC 166
 2230 POGYHLMSEALPLICQALSCGLPEAPKXGMYFG--KQYVTGTAIVYSCSBGYH---LQA 2282
 1054 GNTAH--WSTRPPIQRIPIPCLPPIANGDFISTRENPHYSVVTYRCNLSGRKRVFEL 1112
 2283 GAETAECL---DTGLMS-----NNVVP-----OCV 2306
 1113 VGEPSIYCTSNDOQVGIWSGAPPOCTIPNKTTPNVEKGLVSDNRSLSFLNEVVEFRQ 1172
 2307 P-----VTC-----PDVSSIS-----VHGRWRLIFETQYFOALMLICD 2342
 1173 PGFVPMKPRVVKQALNKWEPELPSGRVCORPEPILHGHNTSHONFSGQEVFPYSC 1232
 2343 PGYTYTGQVIRICOANGKWSLGDSTPTCRIISC--GELPIPNHRICTGLSV-YGATA 2397
 1233 PGYDLRGAASLHCTPQDWS--PEAPRCVAKSCDDFLGQ---PHGVLPPLNLQGAKV 1287
 2398 IFGNSGYTLVGSRYRCMANG--LMSGEVCLACHCTPREIYVNGHNGE--NYSY 2451
 1288 SFVCDSEFRLKGSVSHCVLVGKRSLLMNSVPPVCEHIFCNPRAILNGRHTGPSGDIPI 1347
 2452 RGSVYQCN-----AGFRLIGMSVRIQODNH-----WSGKTPFC--VPITCGH----- 2493
 1348 GKELSTYCDHPDRKMTFNLIGESTIRCTSDPHNGVWSSPARCELSVAAGHCKTBEQF 1407
 2494 ---PGNVNGLTQGNQFNINDVVKFYCNGYMEGAARSQCLASGQSMPLPTCRLIINC 2549
 1408 PFASPTPIINDF---EFVGTSLNVECRPGYEGKMFIS--CLENLWSSVEEDNCRKRSK 1462
 2550 TDPGHQENSRYQVHASBPHRFSGTIVSYVCHNGFYLTGPVLSQDGD---TMDRPRQ 2606
 1463 GPPEPFGNVHINTD---TQFSTVNSCNSGFRILIGPSTTCLVSGNNVTDKCAPT 1518
 2607 CLVSCGHPSPHSGWSDSYT-----VGAVVRYSC-----IGKRLVGN 2647

Db 1519 CEIISG---EPPEITISNGDFYNNRTSFHNGTVVYQCHTGDPDEQLFEIVGERISY-- 1572

Qy 2648 STRMCGLDGHWTSGLPHCSGTSGVCGDPGIPAHGIRLDS---FDGTMVRSCECHV 2704

Db 1573 CTSDDQGVWSSPPFCISITN--KCTAPEV-ENALVPGRRSFLTEIRFCQGFV 1629

Qy 2705 LRGSSEPTQANGSWSSQPECGVISGCPNPTPSNAVFS--DGLVFSSSIYECREY 2762

Db 1630 MVGSHVYQCCNTRMGWPKLPKCSRV-CQPPREILHGHTLSHQNFSPQGEVFSCEPSY 1688

Qy 2763 YATGLSRHCSVNGTWGSDPECLVINCQD--PGIPANGRLGNDPFRYKNTVYQCYPGY 2820

Db 1689 DLRAAALHCTPQDWSPEARCTVKSQDPLQGLPHGRVLLPLNLQLGAKVFSVCDPGR 1748

Qy 2821 MMESHVVS--VLSTCKRTWNGTKPVCKALMCKPPLIPNGKVVS--DEMMGSVTVY 2875

Db 1749 RUKRSASHCVLAKMK-ALMNNSSVPCVCEQIFCPNPPILNGRHTGTFPGDIPYGEISYA 1807

Qy 2876 C-----LEGQLSLPAVFTCEGNSWTGELPQC 2903

Db 1808 CDTHPDRGMTFNILIG-ESSIRCTSDPQNGVWSSPAPRC 1845

RESULT 9
5472939-10
; Patent No. 5472939
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; MONG,
; MINNIE W.; GASSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138,825
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588,128
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:10:
; LENGTH: 2006
5472939-10

Query Match 7.5%; Score 1267.5; DB 6; Length 1847;
Best Local Similarity 21.7%; Pred. No. 7.6e-89;
Matches 503; Conservative 253; Mismatches 768; Indels 795; Gaps 103;

Qy 798 CEEPEVPA--VSIRKGLQGVGDTLTFSCFPGYRLGRTARIC-LGRRRLMSSPLRPG 853

Db 109 CRNPBPVNMVAVIKIOP--GSQIKYCTKGYRLIGSSSATCIISGDTIVINDTTPIC 166

Qy 854 -VAECANSVTGTGLLSPNFPVNNNNECIYISITOPKGIQLKARAFELSEGDVLYK 912

Db 167 DRIPCGLPITTINGDFTSTN--RENPHVGSVVTYRCNPSG--GRKVFEL-VGSP-SI 218

Qy 913 YDNNNSARLLGVFSHSEMMGVTLNSTSSLMWDFITDAENISKGELAFPSFELIKCD 972

Db 219 YCTSDNDQ--VGIMSGPAPQCIIIPN-----KCTP 245

Qy 973 PGTPEKGYKVHDEGHF--AGSSVSFSCDPQYSLRGSEBLLCLSGEBRTVDRPLPTCVAR 1030

Db 246 PNVEN-GILVSDNRSLPSLNEVEFRQPFVWKGRRVYCAQANK--MEPELPSGSRVC 302

Qy 1031 GGTVRGEVSGQVLSPEGPAPYEHNLCIWTIEAAGCTTGLHFLVDETEVDVLRIMDG 1090

Db 303 -----QPP-----DVLT----- 310

Qy 1091 PVSSGVLLKELSGPALPKOLHSTFNSVLIQFTDFTSKQGFALQPSVSTATSCNDPGR 1150

Db 311 -----A 311

Qy 1151 QNGSRSGDSWEAGDSTVFQCDPQVALQSGAEISCVKIENRFWMQSPPTCIAPCGDLTG 1210

Db 312 ERTORDKDNFSPQGEVFSCEPGYDLRGAASMRCTPOGD--WSPAPFCEVASCDDFMG 368

Qy 1211 P--SGVILSPNYPPEPFPKCEKDMKTVPDVLVFNFIENLEPVDLHIYDGDLSLS 1268

Db 369 QLLNGRVLP-----VNLQAKVDVCD----- 392

Qy 1269 PLIGSFYQQLPGRLESSNSLFLA-----FRSDASVSNAGFVIDYTENPRESCEPDGSI 1323

Db 393 -----EGFQKNG--SSASYCVLAGMESLWMSVPCVQIF-----CSPPIY 432

Qy 1324 KNGTRVGSULK-----GSSVTTYCH-----GGVEVGTSTLSCILGPDCKPVWNNRPVC 1374

Db 433 PNGRHTGKLEVPFPGKAVNYTCDHPDRGTSFDLIGESTRICTSDPQNGVWSSPAPRC 492

Qy 1375 TAPCGQVYVGSQVVLSPVYQNYTSQICLYFVTVPKDYVVGQAPFHTALNDVVEVH 1434

Db 493 -----GIL----- 495

Qy 1435 DGHSG--HSRLLSLSGHTGESLPATSNOVLIRFSAKGLAPARGF-----HFVYQAV 1486

Db 496 -GHQAPDHFLLPAKLTQTNASDPIGTS--LKYCEPVEYGRPFSTICLDNLVWMS- 549

Qy 1487 PR--TSATQSSVPEPRYK-RLGSDFSVGAIVRECNAGVYLOSPREICLPVGAALQ 1543

Db 550 PVDCKRKSCKTPPDPVGMVHVTIDIOVSIRINSCYTGHLISHSAECI-LSGNAAH 608

Qy 1544 MNVASAPTCV-VPQGNLTERGTLSPGPEBYLNSCWKIVPEGAGIQIYVSPFT 1602

Db 609 WTKRPICQRIICGLPPLIANDPISITRENHYS-----VITYC 650

Qy 1603 EONMDSLEVFGADNTVTLMSFSGTTPALINSTNOVLHIFYSDISVSAAGFHEKYT 1662

Db 651 NPGSGRKVFE-----LVGEPS-----LYCTSDN-----DQ 676

Qy 1663 VGLSCPEP-AVPS-----NGVKTGER--YLVNDVYFSQCEPGVALQGHANISM 1709

Db 677 VGIWGPAPDQCIIPKCTPPNVENGLIVSDNLSLNEVEFRQCPGPRRVKQ 736

Qy 1710 PGTVRWVYPPPLCIAQCGGVEEMEGVILSPGPGNVPNMDCMKIALPVFGAHIQF 1769

Db 737 -ALNKMELPSCGRVC----- 752

Qy 1770 LNFSTPEPHNDYIEIRNGPEYETSRMMGRPSGSELPSSLSTSHETTVYFHSQNRPGFK 1829

Db 753 -----QPPPDVL-----HAERTQ----- 766

Qy 1830 LEYQAVLEQCPDPEPFANGIVRGAVVGSVTFECLPGYOLTHPVULTQCHGNRMWD 1889

Db 767 -----DKNF-----SPGSEVFSCEPGYDLRGAASMRCT--PDGWS 802

Qy 1890 HPLKCEVPCGNITTS--NGTVYSPGPPSPYSSQDCWMLITVPIGHVRLNLSLQTE 1947

Db 803 PAAPFCEVKSQDDFMGLNGRVLP-----VNLQAKV----- 837

Qy 1948 PSGDPTTWDPQOQAPRLGVFTSMMAKTVQSSNOVLKFNHRDATGCIFFALFASYP 2007

Db 838 ---DFVCD-EGFQKNG-----SSASYCVL-----AGMESLWMSVP 869

Qy 2008 LTK--CPPTLPFAVAVTEN-BEENIGDIYVRCLP-----GFTLVGNEILTKL-- 2055

Db 870 VCEQIFCSFPYIPNGRHTGKPLEVFPFGKAVNTTCDHPDRGTSFDLIGSTICTSDP 929

Qy 2056 ---GTYIQFEGPPICEV--HCPTNELTSDTGVLISQSYSGSYPOFOTCSMLVVRBDY 2110

Db 930 QNGVW---SSPAPRCGILGHQA-----PD- 952

Qy 2111 NISLVEYFLSEKQYDFEITDQSGQSPLLKALSAGYSAFLITYSSNSYLLRMSG--DH 2169

Db 953 -----HFLAKLTQTNAADPIGTSLKYECPREYGRPFSTCLDNLV---WSSPKD 1002
 Qy 2170 AYNRKGIKIVSAPYGLPAPLHGLTGTSTPGSIHFGCNAGRLVGHSMALCTRH 2229
 Db 1003 VCKRKS-----CTPPDPVNGMVHTIDIQVGRINVSCTTGHRLIGHSSAECILS 1053
 Qy 2230 PGYHLMSEAPLCOALSGLPEAPKNGMVF---KEYVTGTRAVYSCSEGYH---LOA 2282
 Db 1054 GNTAH-WSTKPRICQRIPCGIPPIIANGDPISTRENPHYGSVVTYACNLGSRGRKFEL 1112
 Qy 2283 GAETAECL---DTGLMS-----NRPVP-----OCV 2306
 Db 1113 VGEPSIYCTSDNDQVIGSWGAPQCIIPNKCTPPNVENGLIVSDNRLFLSLEVVDPRCQ 1172
 Qy 2307 P-----VTC-----PDVSSIS-----VEHGRRLIFETQYQFOQMLICD 2342
 Db 1173 PGFVMKPRRVKCOALNKMBEELPSCSRVCOPPEILHGEHTPSHONFSPGOGVFSCE 1232
 Qy 2343 PGYVYTQVYIRCOANGKWSLGDSTPCRIISC---GELPIPNHGRIGTSLV-YGATA 2397
 Db 1233 PGIDLRGAASLHCTPGQDMS--PEAPRCAYKSCDDPLGOL---PHGRVLPPLNLQLGAKV 1287
 Qy 2398 IFSCNSGYTLVGSRVRECMANG--LWGSSEVRCLAGHCGTPEPIVNGHNGE---NYSY 2451
 Db 1288 SFVCDSEFRLKSSVSHCVLGMKSLMNNSVPCENHFCBNPAILNGHGTGTPSGDIPY 1347
 Qy 2452 RGSVYVOCN-----AGRLIGMSVRIICQDDH---WSGKTPEC-VPTTCG----- 2493
 Db 1348 GKEISYCDHPDPGMTFNLIGESTICTSPHNGVWSSPAPRCESLVAGHCKTPEQF 1407
 Qy 2494 -----PGAPVNGLTQGNQFNLDVYKFCVNGCYMABAGARSQCLASGQMDLPTCRLINC 2549
 Db 1408 PPSFTPIPIINDF---EPVGTSLNTECRPGYFKMFSIS-CLENLWSSVEDNCRKXSC 1462
 Qy 2550 TDPGHQNSVROYVHASGPHRFSGTIVSYRCNHGFYLLGTVPVLSQCGDG---TWDRPRQ 2606
 Db 1463 GPRPEPFGMVHTID---TQFSTVNYSCNEGFRLIGSPSTCLVSGNNVTMDKAPL 1518
 Qy 2607 CLVSCGHPSGPPHSQMSGDSYT-----VGAVVRYSC-----IGKRTLVGN 2647
 Db 1519 CEIISC---EPPTLNGGFYSNNRTSPFHNGVTVTVOCHTGPDEQLFELVGRSITY-- 1572
 Qy 2648 STNMGGLDGHMTSLPHCSGTVGVCGDPGIPAHGILGDS---FDBGVWFSCEAGHV 2704
 Db 1573 CTSKDDGVWSSPPRCISTN--KCTAPRY-ENALVPGNRFSFSLTEILRFRCQGFV 1629
 Qy 2705 LRGSSERTCOANGSWSGOPECVISCNPPTPSNARVFS--DGLVFSSSIYECREGY 2762
 Db 1630 MVGSHTVQCTNGRMGKLPKHSRV-CQPPPEILHGHHTLSHONFSPGOGEVFSCEPSY 1688
 Qy 2763 YATGLLSRHCSVNGTWSDECLVINGCD--PEIPANGRLGNDPFRYNTVYQCVPGY 2820
 Db 1689 DLGGAALHCTPGQDMSPEAPRCVKSDDPLGOLPHGRVLLPPLNLQLGAKVSFVCEGF 1748
 Qy 2821 MMESHRS--VLSTKRTNNGTKPVCKALMKCRPPLIPNGXVGS---DFMWSSTIYA 2875
 Db 1749 RLKGRASHCVLAKMK-ALMWSVAVCEQIFCPNPPAILNGRHTGTPEGDIYKELISTYA 1807
 Qy 2876 C-----LEGYQLSLPAVPTCEGNGSGMTGELPQC 2903
 Db 1808 CDTHPDGKMTFNLIG-ESSIRCTSDPGONGWSSPAPRC 1845

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/588,128
 ; FILING DATE: 24-SEP-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 412,745
 ; FILING DATE: 26-SEP-1989
 ; APPLICATION NUMBER: 332,865
 ; FILING DATE: 03-APR-1989
 ; APPLICATION NUMBER: 176,532
 ; FILING DATE: 01-APR-1988
 ; SEQ ID NO: 6
 ; LENGTH: 1466
 5256642-6
 Query Match 6.7%; Score 1133; DB 6; Length 1466;
 Best Local Similarity 23.4%; Pred. No. 1,6e-78;
 Matches 402; Conservative 197; Mismatches 620; Indels 498; Gaps 77;
 Qy 1494 CSSVPEPRYVK-PLGSPFSVGAIVRFECNNGVALQSGPBIETCLVPFGALAQMNVSAPTCV 1552
 Db 52 CKTPDPVNGMVHTIDIQVGRINVSCTTGHRLIGHSSAECT-LSGNAHMTKPRICQ 110
 Qy 1553 -VPCGNLTERRGTILSPGPEBYLNSLNCWKIIVPEGAGIQIOWVSFVTEQNMDSLEV 1611
 Db 111 RIFCGLPPIIANGDPISTRENPHYGS-----VVTYRCNPGSGGRKV 152
 Qy 1612 PDGADNTVTLGSGFSGTIVALLNSTSNQLYLHFYSDISVSAAGFHELYTGLSCPEP 1671
 Db 153 FE-----LVGERS-----LYCTSDN-----DOVGLWSGAP 178
 Qy 1672 --AVPS-----NGVKTER---YLNDVSPCEPGYALQOHAHISCMPTGRVEMY 1718
 Db 179 QCIIPKCTPPNVENGLIVSDNLSFSLNVEFERCOPGVMMKPRRVKQ--ALNKMP 236
 Qy 1719 PPLICAOCCGTVENEGVILSBPGFNGYPSNMDCSWKIALPVGFAHIOFLNFSTEPNH 1778
 Db 237 ELPSCSRVC-----QPP 249
 Qy 1779 DYIETRNGPYETSRMGRFSGSELPSLSLSTSHETTVYFHSDSQNRPGKLEYQAYELQ 1838
 Db 250 DVL-----HAERTOR----- 259
 Qy 1839 ECPDPEPFANGIVRGAGVNVGOSVTFECLPGYOLTGHPVLTQCHTRNMDHPLPKCEVP 1898
 Db 260 ---DKXNF-----SGQEVYFSCGTYLDRGAASRCT--PGQDMSPAAPTEBVK 304
 Qy 1899 CGGNITSS--NGTVYSPGFSPSPYSSQDCVWMLITVPIGHGVRLNLSLQTEPSGDFITW 1956
 Db 305 SCDDFMQQLNGRVLP-----VNLQLGAKV-----DFVCD- 335
 Qy 1957 DGPQCTAPRLGVFTSRMAKKTVOSSNQVLKTHRDAATGCIPIALASAVPLTK--CPR 2013
 Db 336 EGFQLKNG-----SSASYCVL-----AGMSLWMSVAVPEQILRCP 371
 Qy 2014 PTLIPNAEVVTEN-EEFNIGDIVRYRCLP-----GFTLVGNELTQKL-----GTLYQF 2061
 Db 372 PVYIPNHRHTGKPLEVFPFKAVVNTIDPPRDGTSFDLIGESTIRCTSPQNGWV--- 428
 Qy 2062 EGPPICEV--HCPTELLTDSGTGVLISQSPGSPQFOTCSMWLVREVPDYNISLTYEYF 2119
 Db 429 SSPAPRCGLIGHQA-----PD-----HF 447
 Qy 2120 LSEKQYDEFEIIPGSGQSPFLKALSGNYAPLIVTSSNSVYLRMS-DHAYNRKGFKI 2178
 Db 448 LFAKLKTQTNASDPFISTSLKYECPREYGRPFSTCLDNLV---WSSPDVCKRKS--- 501
 Qy 2179 RYSAFVSLRAPLHGLTGTSTPGSIHFGCNAGRLVGHSMALCTRHPOGYHLMSE 2238
 Db 502 -----CTPPDPVNGMVHTIDIQVGRINVSCTTGHRLIGHSSAECILSNTAH-WST 554
 Qy 2239 AIPLOALSGLPEAPKNGMVF---KEYVTGTRAVYSCSEGYH---LOAGAETAECL 2291
 Db 555 KPRICQRIPCGIPPIIANGDPISTRENPHYGSVVTYRCNLGSRGRKFELVGEPSIYCT 614

RESULT 10
 5256642-6
 ; Patent No. 5256642
 ; APPLICANT: PEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
 ; MINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
 ; H.; MAKIDES, SAVVAS; MARSH, HENRY C. JR.
 ; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
 ; RECEPTOR 1 (CR1) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
 ; USE THEREOF
 ; NUMBER OF SEQUENCES: 30

```

QY 2292 ----DTGLMS-----NRNVP-----QCV-----2307
DB 615 SNUDDVGIMSGAPPOCIIPNKCTPPNVNGLIVSNRSLFSLNEVERECQPFVWKCP 674
QY 2308 -VTC-----PDVSI-----VEGRMLIFETQYFOQALMLICDPGYTGOR 2351
DB 675 RYKCALNKMEBELSCSRVCCPPEILHGEHTPSHQDNFSBGQVFCSEGYDLGAA 734
QY 2352 VJRCQANGKSLGSDTPTCRITSC-----GELIPBNHRIGHTLSV-YATALIFECNSGYT 2406
DB 735 SLHCTPRDMS--PAPRCAYKSCDPLGOL--PHGRVLPPLMLQGLAKVSPVCDGFR 789
QY 2407 LVGSFVRECMANG--LWSGSEVRCLAGCHGPBEIVANGHNGE--NYSYRGSVYQCN 2460
DB 790 LKGSSEVSHCVLGMKSLMNNNSVPVCEHIFCPNPAIILNGRHTGTPSGDIPRGKISTCD 849
QY 2461 -----AGPRLIGMSVRIICODH-----WSGKTPEC-VPTTCG-----PSNPV 2498
DB 850 PHPDRGMTFNLIESTRICTSDPHGNGVWSSPAPRCESLVRAGHCKTPEQPFASPTIPI 909
QY 2499 NGLTQGNQFNLDVVKFCVCPNGYMAEGARSCCLASGQMSDMLPTCRILNCTDGHQENS 2558
DB 910 NDF-----EFYGTSLNTECRBGYFGKMSIS-CLENLWSSVEVENCRRKSCGPPPEPFG 964
QY 2559 VROVHASGPPHRSFGTTVSYRCHGFYLLGTPVLSGQDG--TWDRPPOCLLVSCGHP 2615
DB 965 MVHINTD-----TOGSTVYNWSCNEGFRILIGSPITCLVSGNNVTWMDKAPICEIISC-- 1017
QY 2616 GSPHSQMSGDSY-----VGAVVYSC-----IGKRLVGNSTRMCLDG 2656
DB 1018 -EPPPTISNDIFYNNKRTSPFNGTVVYQCHTGPDSQLFELVGERISY--CTSKDQVQ 1074
QY 2657 HMTGSLPHCSGTAVGCGDPGIPAHGIRLGS--FDPGTVMRFSCEAGHVLGSSERTC 2713
DB 1075 VMSGPPRCISTN--KCTAPEV-ENALRVPGNRFFSLTEIIRROCPGVVMGSHVQC 1131
QY 2714 QANGSMGSGOECGAVICGNPRTSNARVFS--DGLVFSSTIYECREBYATGLISRH 2771
DB 1132 QTNRMWPKLRPHGRV-CQPPPEILHGEHTLSHQDNFSPQGEVYFSCPEPYDLRGAASH 1190
QY 2772 CSVNGTSGDPECLVINCGD--PGIPANGRLGNDPRYKTVYOCVPYMMESHRSV- 2828
DB 1191 CTGPGDMSPEAPRTVASCDDPLGQLPHGRVLPPLMLQGLAKVSPVCDGFRILGKSASH 1250
QY 2829 -VLSCTDRTWNGTKVYCKALMKCPPLINGKAVGS--DFWKGSSVYAC-----2876
DB 1251 CVLAGMK-ALMNSSVPCQEIFCPNPPAILNGRHTGTPFGDIPYKEISYACOTHPDRGM 1309
QY 2877 ---LEGVQLSLPAVFTCEGNSWTGELPOC--FPVPCGDPGVPSGRREDRGFSY--R 2927
DB 1310 TPNLIG-ESSIRCTSDPDQNGVWSSPAPRCESLVPACPHPKIONGHYIGHVLSYLPG 1368
QY 2928 SSVFSCHPPLVLVGSFRRFCOSDGTWSTGTPSCIDPTLTTCADPVPFGIQQN--SO 2984
DB 1369 MTSYTCDDPGYLLVVGKGFICTDQGLWSQDLHCKE--VNGSFP-LFMNGISKELEMK 1424
QY 2985 GYQVGSVTLFRCKQKYLLOSSTRTCLPMLTWSGTPP 3021
DB 1425 VHYGYVTLKCEDGYTLGSPWSQCADDRWD--PP 1459

```

```

RESULT 11
5472939-6
; Patent No. 5472939
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LOYD B.; WONG,
; WINNIE W.; CASON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138,825

```

```

; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588,128
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO: 6
; LENGTH: 1466
5472939-6
Query Match 6.7%; Score 1133; DB 6; Length 1466;
Best Local Similarity 23.4%; Pred. No. 1.6e-78;
Matches 402; Conservative 197; Mismatches 620; Indels 498; Gaps 77;

QY 1494 CSSEVPEPRYCK-RLGSDFSVGALVIFECNSGYALOGSPREIECLPVPGLAONNVSAPTCV 1552
DB 52 CKTPDPVUNGMVHVITDIDVGSRIYSCITGHRLLIGHSSAECT-LSGNAAHMSTKPPICQ 110
QY 1553 -VPCGNLTERGTLISPGFPPEPYLNSLNCWKIIVPEGAGIQIOWSVFTEQNDSELEV 1611
DB 111 RIPCGLPPIANGDFISTRENPHYGS-----VVTYRCRPSGGRKV 152
QY 1612 FPGADNTVMGSPFGSTVYPALNSTSNQYLHFYSDISVSAAGHLEKYVGLSSCEP 1671
DB 153 FE-----LVGEPs-----IYCTSDND-----DQVIGSGAP 178
QY 1672 -AVPS-----NGVATGER--YLNDVYSPQCEPFGYALOGHAH:SCMPGTARRNNY 1718
DB 179 QCIIPKCTPPRVENEGILVSDNRSLFSLNEVERECQGFVWKGRYKCO--ALNKWEP 236
QY 1719 PPLCLAQCGGTVEEMEGVILSPGFPNGVPSNMDCKIALPVGGAHIOFLNFSTERNH 1778
DB 237 ELFSCSRVC----- 249
QY 1779 DYIEIRNGPYETSRMWRGFSGSELPSLSLSTSHETTVYFHSDHQNRGFLVQAYELQ 1838
DB 250 DVL-----HAEKTR----- 259
QY 1839 ECPDPEPPANGIVRGAGYNGQSVTFECLPGYOLTGHPVLTCQHTNENMDHPKCEVP 1898
DB 260 ---DKDNF-----SPQGEVYFSCPEPYDLRGAASMKCT--PQSDWSPAPPTCEVK 304
QY 1899 CGGNTSS--NGTYSPFPSPYSSQDCWLITVPYGHVRLMLSLQTEPSEGDFITW 1956
DB 305 SCDDMGQLNGRYLFP-----VNLQGLAKV-----DFVCD- 335
QY 1957 DGPOQTARLGVFTRSMAKKTVOSSNOVLKPFHRDATGIFAIAPSAVPLTK--CP 2013
DB 336 EGFOLKG-----SSASYCVL-----AGMESLMNSVPCQILCP 371
QY 2014 PTLIPNAEVTEEN-EEFNIGDIYRYCLP-----GFTLVGNELLTKL-----GYLQF 2061
DB 372 PPVIPNGRHTGKPLEVPPFKAVNYTCDPHDRGTSFDLIGESTRICTSDPDQNGWV-- 428
QY 2062 EGPPICEV--HCPTELNDTSTGIVLISQSYPGSPQFQCSMLVRPEPDYNISLTYEYF 2119
DB 429 SSPAPRCGILGHCA-----PD-----HF 447
QY 2120 LSEKQYDFEFLFDPSQSPPLKALSGNYGAPLIVTSSNSVYLRWMS-DHAAYNRKQFKI 2178
DB 448 LFAKLKTQTNASDPPIGTSLYKBRPPRYGRPSITCLDLV--WSPPDVCKRKS-- 501
QY 2179 RYSAIPYCSLPAPPLHGTLIQTSTQPGGSHFGCNAGYRVLVGSMAICTRHPOGYHLMSE 2238
DB 502 -----CKTPDPVUNGMVHVITDIDVGSRIYSCITGHRLLIGHSSAECILSGNTAH-WST 554
QY 2239 AIPLOALSGLPPAPKNGWVFG--KEYTVGKAVYSCSEGVH--LOAGNEATNECL 2291
DB 555 KPPIQARIPOGLPPTIANGDFISTRENPHYGSVVTYRCNLGSRGRKRVFELVGPSPSIYCT 614

```



```

01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1
```

```

; APPLICATION NUMBER: US/08/588,128
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:5
; LENGTH: 1537
; 5256642-5
;
Query Match      6.7%; Score 1133; DB 6; Length 1537;
Best Local Similarity 23.4%; Pred. No. 1,78-78;
Matches 402; Conservative 197; Mismatches 620; Indels 498; Gaps 77;

1494 CSSVPEPRYK-RIGSDPSVGAIVREFCNSGYALQSGPEIECLPVPALQAMNVSAPTCV 1552
52 CKTPDPDVGNGVHVITDIQVSRINYSCTTGHRILGHSSACI-LSGNAAHMSTKPRICQ 110
1553 -VPGGMLTRRGITLSPGFPEPLNLNCWKIVBEGAGIQOVVSPFTQEWMDLSEV 1611
111 RIPGGLPPTIANGFISTNRENFYGS-----VVTIRCNPGSGRKV 152
1612 PDGADNVTVMIGSPSGTTVPALLNSTSNQULHFESDISVASAGPHLEVKTGSLSCPEP 1671
153 FE-----LVGERS-----LYCTSDN-----DQVGIWSPAP 178
1672 --AVPS-----NGVKTER---YLVDVVSFOCEPGYALQGHAIHSCMPGTVRMNY 1718
179 QCIIIPNKCTPENVENGILVSDNRSLFSLNEVEPRCQGFVWKQPRRVKCO--ALANKEP 236
1719 PPELICQCGGTVBEMGCVLLSPFPNGNYSNMDCSKIALPVGFGAHIQFLNSTEPNH 1778
237 ELFPCSRVC-----CPP 249
1779 DYIEIRNGPYETSHMGRFSGSELPSLLSTSHETTVPHSDHSQNRPGFLLEYQAYELQ 1838
250 DVL-----HAERQR----- 259
1839 ECPDPEPPANGIVRAGAYNQGVSTFECLEPGYOLTHPVLTCQHTNRNMDHLPKCEVP 1898
260 ---DKNDF-----SPQGEVFYSCBPGYDLRGAASMKCT--POGDWSPAAPTCBEVK 304
1899 CGGNITSS--NGTVYSPGPSPSPSSQDCWLIIVPIIGHGRVRLNLSLQTEPSCDPIITW 1956
305 SCDFPFGQNLGKRVLP-----VNLQGAKV-----DFVCD 335
1957 DGPQOTAPRLGVFTRSMNAKTVQSSNQVLLKFRDADATGGIFAIASAVPLTK---CP 2013
336 EGFPGLK-----SSAYCYVL-----AGMELMWSVAVCEQICPS 371
2014 PTLIPNAEVTEN-EEFNIGDIYRIKLP-----GTLVGNELITKCL-----GYLQF 2061
372 PPVILPNRHHGKPLEVFPFGKVAVYTCDPPHDSRTSDLLIGESTIRCTSPDQNGW--- 428
2062 EGPPIICEV--HCPNTNELDSTSTVILLISQSPGVPQFQCSMWLREVPDNLISLVEYF 2119
429 SSPAPRCGIIIGHCA-----PD-----HF 447
2120 LSEKQYDFEFLFDGSPGSPPLKALSGNYSAPLIVTSSNSVYLWMS--DHAYNRKGFKI 2178
448 LFAALKQTMAASDPPIGSLKYECRPEYGRPFSTICLDMLV--WSSPKDVCKRKS--- 501
2179 RYSAPYCSLPAPLHGFTILGQTSIQPGCSIHFGCNAGYRLVGHSMALCTRHPOGYHLMSE 2238
502 -----CKTPDPDVGNGVHVITDIQVSRINYSCTTGHRILGHSSACIISGNTAH-WST 554
2239 AIPICGLASGLPEAPKNGWVFG---KEYVTGTAAVYSCEGYH---LQAGATAECL 2291
555 KPPIICQIIPCGLPPTIANGFISTNRENFYGSVYVTRCNLGSNGKRVFELVEGSPSYICT 614

```

```

QY 2292 ----DTGLMS-----NRNVP-----OCVP-----2307
Db 615 SNDQVGMGAPACQIIPNKCTPBNVENGILVSDNLSFLSNEVEFRQCPVWKGPR 674
QY 2208 -VTC-----PDVSSIS-----VEKRMILFTQYQFQAQMLICDPGYTGTQR 2251
Db 675 RVKQALNKWPELPSQSVQCPPELHGHHTPSHDNPSPOGEVFSCEPGDILGAA 724
QY 2352 VIRCOANGKMSLGDSTPTCRIISG---GELPIPNHRIIGTSLV-YGATAISCSNGYT 2406
Db 735 SLHCTPRGDMSS--PEARCAVKSCDLFLGQI--PHGRVLPFLNLQICAVYFVCDGFR 789
QY 2407 LVSGRVEECNANG--LMSGEVACLAGHCSTPEPIVNGHINER--NYSRGSVVYQCN 2460
Db 790 LKGSVSHCVLVGMRSIMNNSVPVCEHIFCPNPPALINGHHTGTPSGDIDYKBIYSTCD 849
QY 2461 -----AGFRLIGMSVAVICQODHH-----MSGKTPFC-VPTTCGH-----PCNPV 2498
Db 850 PHPDGKMTFNLIGESTIRCTSDPHGNGVSSPAPRCLSVRAGHCKTPEQFPAPSPITPI 909
QY 2499 NGLTQGNQFNLNDVVKFVCNPNGYMAEGAAQSCLASGQMSDMLPTCRIIINCTDPGHQENS 2558
Db 910 NDF-----EPVUGSILNTECRPGYTGKMFIS-CLENLVSSVEDNCRKSCGPPPEFNG 964
QY 2559 VROVHASGPHRFSGTIVSYRCHNHFYLLCTPVLSCQDQ--TWDRPRQCLLVSCGP 2615
Db 965 MVHINTD-----TOFGSTVNSCNEGFRILGSPSTTCLVSGNNVTDKACIDCEIISC--- 1017
QY 2616 GSPHSGMSGDST-----VCAVVRYS-----IGKRLVANSIRMGCLDG 2656
Db 1018 -BEPPTISNDPFSNNRTSFHNGTVVYQCHTGPBGOLFELVGERIY--CTSKDQVQ 1074
QY 2657 HMTGSLPHCSGTSVGVCGDPIPAHGIRLGD--FDPGTVMRFSCAGHVLGSSERTC 2713
Db 1075 VMSSPPRCISTN--KCTABEV-ENALRVGNMSFELTEIIRFCQPGVVMGSHTVQC 1131
QY 2714 QANGSWSGQPECGVISCNPGTPSNARVFS--DGLVFSSSVYECREGYVATGLLSRH 2771
Db 1132 QITGRMGPRLPHCSRV-CQPPPELHGHHTLSHDNPSPOGEVFSCEPSYDLRGAALH 1190
QY 2772 CSVNGTMTGSDPRECLVINCDD--PGIPANGLRLGNDRYKKTYYQCVPRYMESHRVS 2828
Db 1191 CTRQGDMSPEAPRCVTVASCDPLGQDPHGRVLLPLNLQIGAKVSPCDBEFRLKGSASH 1250
QY 2829 -VLCTDRMTNGTKPVCKALMKCPRELIPIHGKVVGS--DFMWGSSVTVAC-----2876
Db 1251 CVLAGMK-ALMNSGVPCEQIFCPNPPALINGHHTGTPFDIYKKEIISACDTHPDRCM 1309
QY 2877 ---LEGQSLPRAVFTCEGNSWTGELPQC--PVPFCGDPGVPSRGRREDRGFSY--R 2927
Db 1310 TFWLIG-ESSIRCTSDPOGNGVSSPAPRCELSVPAACPHPKIONGHYIGHVSLLPQ 1368
QY 2928 SSVFSCHPPLVTVGSPRRCQSDGTWMSGTQPSCIDPTLTTCADPGVPOGCIQNN---SQ 2984
Db 1369 MTLSTYCDPEYLVGKGFICTQGLMSQLDHYCKE--VNCSPF-LFMNGISKELEMKK 1424
QY 2985 GXVGSFVLFRCCQGYLLQSTTRTCLPMLTWSGTP 3021
Db 1425 VHYGDVYTLKCGEDYTLBSSPWSQCADRMD--FP 1459

```

```

RESULT 13
5472939-5
; Patent No.: 5472939
; APPLICANT: PEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; MINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/06/138, 825
; FILING DATE: 19-OCT-1993

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588,128
; FILING DATE: 24-SEP-1980
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:5
; LENGTH: 1537
5472939-5
Query Match 6.7%; Score 1133; DB 6; Length 1537;
Best Local Similarity 23.4%; Pred. No. 1,76-78;
Matches 402; Conservative 197; Mismatches 620; Indels 498; Gaps 77;
QY 1494 CSSVPEPRYRK-RLGDSFSGAIVRECNAGVALQSGPEIECLPVPGALOMNVNAPTCV 1552
Db 52 CKTPDPVNGMVHVTIDIVGSRINVSCTTGHRLIGHSSAECL-LSGNAAHMSTKPEPICQ 110
QY 1553 -VPCGNLTERGTLSPGPEPIYLSNLCWKIYVBEAGIQIYVSVTEQNMDSLEY 1611
Db 111 RIPCGLPTIANGDFISTNRENPHYS-----VVTYRCNPGSGGRVY 152
QY 1612 FCGADNTVMLGSGGTIVPALLNSTNQYLHFYSDISVSAAGPHLEKTYGLASCPEP 1671
Db 153 FE-----LVGEPS-----IYCTSDN-----DQYINGSPAP 178
QY 1672 --AVPS-----NGVKTGER--YLVDVVSFOCEPGVALQGHANISCMGTVARMY 1718
Db 179 OCILIPKCTPBNVENGILVSDNLSFLSNEVEFRQCPVWKGPRVACQ--ALNKKWP 236
QY 1719 PRELCAOCCGVEEMEGVILSPGFRGNVPSMDCSKIALVVGCAHIQFLNFTSRPH 1778
Db 237 ELPSGSRVC-----QPPP 249
QY 1779 DYIEIRNGPEYETSRMNGRPSGSELSLSTSHETTVVPHSDQGNRPGFKLEYQAYELQ 1838
Db 250 DVL-----HARTOR----- 259
QY 1839 ECPDPEPPANGIVRAGVNGSVTFECLPGYQLTGHVPLTQCHGTNRNWDHPLKPEVP 1898
Db 260 ---DKDNF-----SPQGEVFSCEPGYDLRGAASMRCT--PQGMSSPAAFPCEVK 304
QY 1899 CGGNITSS--NGTVSPGPRSPSYSSQDCVMILITPIGIGVRLNLSLQTESGDPITIT 1956
Db 305 SCDDFMGOLLNGRVLFP-----VNLQIGAKV-----DFVCD- 335
QY 1957 DGPQOTAPRLGVFTSRMAKKTVOSSSNQVLLKFRHDAATGIGIFALAFSAVPLTK---CP 2013
Db 336 BGFOLKG-----SSASYCVL-----AGMSLMSNSVPCQIFCFS 371
QY 2014 PTLTNAEAVTEN-EEFNIGDIVRYCLP-----GFTLVGNELITCKL-----GTYYLF 2061
Db 372 PVIIPNGHHTGKPLRVFPFGKAVNYTCDBHPDRGTSFDLIGSTIRCTSDPGNGVW--- 428
QY 2062 BGPPIICEV--HCPINELLTDSGVILSOSYVGSIPQOTCSWLVRVEDNNYISLVEYF 2119
Db 429 SSPAPRCGLIGHCQA-----PD-----HF 447
QY 2120 LSEKQYDEFELFDGSGOSPLIKALSGVNSAPLIVTSSNSVYLLRMS--DHAYNRKGFKI 2178
Db 448 LFAKTKGTQNMADFDIGTSLKKECPREYGRFSTICTDNLV---WSSPKVDCKRKS--- 501
QY 2179 RYSAPYCSLPRAHLGHFLIGTSTQPGSIHFGCNAGRYLVGHSNAICTRHPQGYHLMSE 2238
Db 502 -----CKTPDPVNGMVHVTIDIVGSRINVSCTTGHRLIGHSSAECLISGNTAH-WST 554
QY 2239 AIPLOALSCGLPEAPKNGMVG---KEYTGTAKVYSCSEYH---LQAAEATABCL 2291
Db 555 KPPICORIPCGLPPTIANGDFISTNRENPHYSVVTYRCNLSGRKRVFELVGEPSIYCT 614

```

QY 2292 ----DTGLMS-----NRNVP-----OCVP----- 2307
DB 615 SNDDQVIMSGPARQCIIPNKCTPRNVENGLVSDNLSLFEINVEFRQCPGVMGKPR 674
QY 2308 -VTC-----PDVSSIS-----VEHGMRLIFETQYQFOALMLCDPGYYTGOR 2351
DB 675 RVKQALNKMEPELPSCSRVCQPPPEILHGHNTSHODNFSPQGEVYSCPGYDLRGA 734
QY 2352 VIRQANQKXSLGDTPTCRISC-----GELPIPNHRICTLSV-YGATAPESCNSGYT 2406
DB 735 SLHCTPRGDSW--PEARCAVAKSCDPLGOL--PHGRVLEPLMLQGAKEYFCYCDGCFR 789
QY 2407 LVGSRVRECMANG--LMSGEVRECLAGHCTPEPIVNGHNGE--NYSYSGVAVYQCN 2460
DB 790 LKSSSVHCHLVGKRSIMNSVPYCEHIFCPENPAILNGRHTGTPSGDIPYKXISTYCD 849
QY 2461 -----AGFRILGMSVRIQQODHH-----WSGKTPFC-VPIYTCG-----PGNPV 2498
DB 850 PHPRGWTFLIGESTRICTSDPHNGVWSSPAPRCELSVAGHCKTPEGPFASPTLPI 909
QY 2499 NGLTQGNQFNLDVYKFCVCPNGYMAEGAASQCLASQMSMLPTCRLINTDGHQENS 2558
DB 910 NDF-----EPVGTSLNECRPGYFKMFSIS-CLENLWSSVEDNCRKSGPPEPENG 964
QY 2559 VROVHAGPFRFSGTIVSYRCNHFYLGTPVLSCQGDG--TWDRPPQCLLVSCGHP 2615
DB 965 MWHINTD-----TQGSTIVNSCHNGFRLLIGSPITCLVSGNNVTWDKAPICELISC-- 1017
QY 2616 GSPRHQMSGDSTY-----VGAIVRYSC-----IGKRTLVGNSTRMCLDG 2656
DB 1018 -BPPRTISNDPFSNNRTSFPNGTVVTVYQCHTGPDSQLFELVGERSIY--CTSKDQVQ 1074
QY 2657 HMYGSLPHCGSTISVGVCDGECIPAHGIRLQDS--PFGTYMRSCAGHVLKRSSERTC 2713
DB 1075 VMSSPPRCISTN--KCTAPEV-EMAIRVPCNRSFSLTELIRRCOPGFVWGSHTVOC 1131
QY 2714 QANGSMGSOPECVISCNPGTPSNARVVS--DGLVSSSIYECREGVYATGLSRH 2771
DB 1132 QTNRMWPKLPHGRV-CQPPPEILHGHNTSHODNSPQGEVYSCPESTDLGAASLH 1190
QY 2772 CSVNGTGTGSDPEGLVINGD--GAIPIANGLRLGNDRYNTVTVYQCPVGMESHRSV- 2828
DB 1191 CTBQGDMSPEAPRCTVASCDFLQDLPHGRVLLPLNLQGAKEYFCVDEGRFLKGRASH 1250
QY 2829 -VJSTCDRTWNGKIPCKALMKCPRLIPNGKIVGS--DPMGSSVYTAAC----- 2876
DB 1251 CVLAGMK-ALMNSSVPCBOIFCENPPAILNGRHTGTPFGDIPYKXISYACDTHPDRGM 1309
QY 2877 ---LEGYQLSLPAVFTCEGNGSWTGLPQC--FPVFCGDPGVSRGRREDGRFSY--R 2927
DB 1310 TPNLIG-ESSIRCTSDPQNGWSSPAPRCELSVPAACPHPKIQNGHYIGGHVSLYLP 1368
QY 2928 SSVSFSCHPPLVIVGSPRRFCQSDGTWSTGTOPSCIDPDLTTCADPGVPOFGIONN--SQ 2984
DB 1369 MTISYCTDPGLVVGKGFIFCTDQGISQLDHYKE--VNCSPF-LFMNGISLELMKK 1424
QY 2985 GYGVGSTVLPFCQKGYLQGSTTTCPLNLTWSTGTP 3021
DB 1425 VHYGDVYTLKCEBDGYTLBGSWMSQCOADDRMD--PP 1459

RESULT 14
US-09-285-385C-4
Sequence 4, Application US/09285385C
Patent No. 6579702
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S.
APPLICANT: Scott, Ian C.
APPLICANT: Thomas, Christina L.
TITLE OF INVENTION: MAMMALIAN TOLLOID-LIKE GENE AND PROTEIN
FILE REFERENCE: 960296.96111
CURRENT APPLICATION NUMBER: US/09/285,385C
CURRENT FILING DATE: 1999-04-02

PRIOR APPLICATION NUMBER: 60/111873
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 60/080550
PRIOR FILING DATE: 1998-04-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 4
LENGTH: 1012
TYPE: PRT
ORGANISM: mouse
US-09-285-385C-4
Query Match
Best Local Similarity 22.9%; Pred. No. 1,4e-50;
Matches 290; Conservative 170; Mismatches 417; Indels 392; Gaps 48;
4.5%; Score 771; DB 4; Length 1012;
QY 1022 PLPTVAECGGTAVGEVSGVLSPPYAPAYEHNLNCIMTIEAGCTIGLHFLVDTTEV 1081
DB 17 PLPR-----GAVTGDHNSVALDYG-----ALEGEGTQQLHY----- 50
QY 1082 HDVLRITWDGVEGVLKELSGPALPKDHLSTNSVVLQFSTDFTSKQGFALQFSVSTA 1141
DB 51 HD-----PCKAVFWGDI--ALBDDKLFR--IDKAEW-----TK 83
QY 1142 TSCNDPG-----IPNGSRSGDSWEAGDSTVPOCDPG--YALQSGAE 1181
DB 84 PSIDKPEHDGGLLETSARPNNTASIASIQARKKQKDATTLPLNPDTSTNTAKTTSAR 143
QY 1182 I-SCVKIENPFQOPSPPTCIAP--CGDLTGSGLVLS- 1218
DB 144 VRATTSRTERIW--PGVIVPVIGNFTGTQRAIFKQAMHWEKHTCTVFERTDEE 199
QY 1219 -----NY-----PEPYRPGKED-----WKVYSPDYVIA 1243
DB 200 SPVFSYRTCCSGSYGRRGGPOAISIGNCKDFGIVAHELGVVGFWEHTRPRDQH 259
QY 1244 LVFNIFLBERG--YDFLHIYDGRDSLPLLSFGYSQULPGRIBSSNSLFLAFPSDASVS 1301
DB 260 VTIIRRENIQGBEYNFLKMEAGEVS--SLGERTY-----DPSIMHYANRTS-- 304
QY 1302 NAGFVIDYTNPRESCDFPSIRNGTR--VGSDLKLGSSVTVYCHGGEVEGTSLCIL 1359
DB 305 -RGVFLD-TILPRRD-----DNGVAPRTIGQVRSLQGDIAQARKLYK----- 344
QY 1360 GPBGKPVMMNPRVCTAPCCGQYVSGDVLSPYIPONTYSGQICLYFTVY--PKDYV-- 1416
DB 345 -----CPA-CSETLQDITGNFSAGFPNGYSYSHCWIRISVTPREKTIILN 389
QY 1417 FGGAPFPHAL--NDVVEVHDGSHOHSRLSSLSGSHTGESLPLATSNOLVLIKESAKGLA 1474
DB 390 FTSMDLFXKRLCWYDIYELIDGYWRKAPLGRFCGDKIPESL--VSSDSRLMWEFRSSSS 448
QY 1475 PARGFHEVQAVPRTSATQCSSVPEPRYGRRLGSDSVGAIIVFECNSGYALQSGPEIEC 1534
DB 449 LKGGFVAVYEAM----- 460
QY 1535 LPVPGLAQNNSAPTCVPCPGGNLTERRGITISPGPEPYLANSNCWMIIVPEAGIQ 1594
DB 461 -----CGGDITKDAQOISPNVDDYRPSKCEVWIRITVPDGHVG 500
QY 1595 IOVVPVTEON-----WDSLEVPDQADNTVMTLSPSGTTPVPLALNSTNSNOLYHFVSDIS 1650
DB 501 LTRQSFIERHDSCAVDYELIRGPFEDSTLIGHFCYERPEAVKSSANLWKFKVSDGS 560
QY 1651 VSAAGFLLEYKTVGLSCPEPAVPSNGVKTGERVYLVNDVVSF--QCEPGYALOGHAHISC 1708
DB 561 INKAGFAANF-----FKVEDCSMPDHG--GCEGRCVNTLSSYCAQDPGYELAA----- 608
QY 1709 MPGTVRANYPPLCLTAQCGGTVEMEGVLSGFGPNYNSMDCSKIALPVGFAHIQ 1768
DB 609 -----DKTCEVACGCFITKNGTITTSQWPEXETNNKNCWQVAVPQYRISLQ 658
QY 1769 FLNFTSEPN--HDYIEINNGPYETSRMMGRSGSLPSLSLSTSHETTIVYFHSN 1824

```

Db 659 FEAFLENDVCKYDFVEVRSGLSPLDALKHGFCSSEFPEVITQSNNMRVEFSDNTVS 718
QY 1825 RRGFLLEYQAYELQECPPDEPPANGIVAGYNNQGVTFECLPEYOL--TGHPVLTCQH 1882
Db 719 KRGFPAHFPS--DKQECARD---NGCCOECVNTFGSLCRGRNGYRLHENGH---DCK- 769
QY 1883 GTNRWMDHPLPKCEVPCGNTSSNGTVSPGPSPSSSODCWILTVPIGHVRLNLS 1942
Db 770 -----EAGCAVYKISSAGCTILMSRNPWPKYPSRKECTWNISSTIGHRVKITPS 816
QY 1943 LLQTEP-----SGDFTTWDPQOQAPRLGVFTRSMAKTVQSSNOVLKPHRDAATGCI 1998
Db 817 EFEEIOHQECAYDHELETDGLAPILGRFCGSKKDPVVAATSSILFIRYSAS---V 873
QY 1999 FAIPSAVPLTKCPRPTILPAVAEVTEENFNIGDIYARCLPEFTLVGNELTKCKICTY 2058
Db 874 QKRGQAVHSTEC-----GGRLLKA-EVQTKELVSH 902
QY 2059 LQFSGPPICEVHCPTNELLTDSGTVLISQSPGSPFOFQTCVPLVVEPDYNI SLTVEY 2118
Db 903 AQFG-----DNNYFSQARCDWVIVAEQGYVELLFT 934
QY 2119 FLSEKO-----YDEFEIPDPSGQSPPLKALSGNTSAPL-IVTSSNSVYLKRWSSDHAAYNR 2173
Db 935 FEVEEADCGYDFMEAYDGYDSAPRLGRFCG--SGPLEEITYSGDSLMIREFHTDITNK 992
QY 2174 KGFIRYSA 2182
Db 993 KGFHARTIS 1001

```

RESULT 15
US-08-991-408-2

```

Sequence 2, Application US/08991408
Patent No. 6008017
GENERAL INFORMATION:
APPLICANT: ARLETH, ANTHONY J.
APPLICANT: WILLETTTE, ROBERT N.
APPLICANT: ELISHOURBAGY, NABILI A.
APPLICANT: LI, XIAOTONG
TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLOID-LIKE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RAYNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,408
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/034,471
FILING DATE: 02-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: ATG-50038
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

```

```

;
; LENGTH: 1013 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-991-408-2

```

```

Query Match 4.5%; Score 769.5; DB 3; Length 1013;
Best Local Similarity 24.6%; Pred. No. 1,8e-50;
Matches 245; Conservative 131; Mismatches 331; Indels 287; Gaps 33;

```

```

QY 879 NNHECIYSIOTQGGKGIQKARAFELSEGDLKYVDGNNNSARLLGV-FSHSMMEVTIN 937
Db 258 DNHVITLIRENIQPGQ-----EYNFLMEPEGANL--LGERYDPSIMHYARN 302
QY 938 STSSSLMDLFDIDAENTSKGFEIHSSPELICKCEDPG--TPKFGYKVH-DEGFAGSSVSF 995
Db 303 TFSRGWFLDTLLPSR-----DNGIRPALGQRTRLSKGDIQAQ----- 340
QY 996 SCDPGYSLAGSEBELCLSGERTWDRPLPTCYAECGGTYRGVSGCVLSPPGYAPAYENHL 1055
Db 341 -----RKLRYCPA-CGETLQ--BSNGNLSSPPGPNPGPSYT 373
QY 1056 NCIIWTIEAAGCTIGLHFLVPDTEEV---HDVLRIWDGPVSGVYLKELSPALPKDHL 1111
Db 374 HCIMRVSVPGRKIVANFTTMDLYKSSLCWYIYIEVRDGYWRKSPILGRFCQDKLPEVLT 433
QY 1112 STFNSVYLOFSTDFTSKQGFALQPSVSTAVSCNDPGIIPQNGSRGDSWEA-GDISTVQC 1170
Db 434 STDSRMWLEF-----RSSNMVKGKFAVVE- 459
QY 1171 DPGVALQGSABEISCVKIENRFWQSPPTCIAPCGGDLTGSPGVLSPYPPYPRGKXC 1230
Db 460 -----AICGSEIRKNEGQIQSPYPPDYAPRMKC 488
QY 1231 DWKTVSPDYVALVAVENITFNEP---GYDFLHIYDGRDSLPLIGSFYGSQLPRIESS 1286
Db 489 VKKITVSESYHVGGLFQFSEIERHNCAYDYLEVADGTSENSPLGRFCGYKPKPEDIIST 548
QY 1287 SNSLPLARSDASVSNAGFVIDYDTEPNPRESCDPSISNGRVSGLDKLGSVTVY---C 1343
Db 549 SNTLWKKFVSDGTAKAFANFFKE-EDECKAP-----DRGCEGRCLNTLIGSYQCAC 601
QY 1344 HCGYEVEGSTSLCILGPDGKVPWNNPVPCTAPCGQGYVSDGVVLBPVNYQONTSGOI 1403
Db 602 EPGYE-----LQPD-----RSCBAACGGLTKLNGITTTGPKXKPYPPNKN 643
QY 1404 CLYEVTVPKDYVVFQGFAPFHTALNDV-----VEVHDGSHQSRLLSLSGSHTGESLPL 1458
Db 644 CWMQVVAFTQYRISVKEFFFELEGNEVCYDVEIWSGLSSKHLGKFCGALEVPEVI-T 702
QY 1459 ATNSQVLLKFAKGLAPARGF--HFVQVAVRTSATQCSSVPERPYGRLGSDSVGAIY 1516
Db 703 SQFNMRLEFPKSDNTVSKKGFRAHF-----SDKDECS--KONGGQHCVCWNTGSYM 753
QY 1517 RFECNSGYALQGS-----PEIECLPVPGALAOMNVSAPCVVPCGGNLTERRGTILSPGP 1572
Db 754 -CQCGNGFVLHNDKHDCEAEC-----EQKHSPS-----GLITSPWP 791
QY 1573 EPYLNSLNCVKKIYVPEGAGIQIOVVSFTBON---WDSLVEFDGADNTVMTLGSFSGT 1628
Db 792 DKYPSRKECTWEISATPGRIKILAFSEFEIQHOECADVHLEVDGETEKSPILGRLCGN 851
QY 1629 TVPALNSTSNOLYHFPSDISVSAAGHLEKTVGLSSCCEPAVPSNGVCTGERIYVND 1688
Db 852 KIPDELVAIGNKMFREFVSDASVQKGFQATHST----- 885
QY 1689 VVSFQCEPGVALQGHAIHSCMPGTVRRWNYPPPLCIAOCGGTV--EEMEGYILSPGPG- 1745
Db 886 -----ECGGRLLAASKPRRLVYSHAQGD 908
QY 1746 -NYSNMDCSWKIALPVGFGAHIOFLANFSTEP---NHDYIEINNGPYETSRMMGRFSGS 1800

```

Db	909	NNYGVDCMVLVSEKRLBLSFQTFEVEEADCGYDYVELFDGLDSTAVGLGRFGS	968
Qy	1801	ELPSSLSTSHETTVFHSDBSONRPGFKLEYOA	1834
Db	969	GPPEIYISIGDSVLHFDHDDTINKKGFHIRYKS	1002

Search completed: October 18, 2004, 10:26:21
Job time : 74 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 18, 2004, 09:55:06 ; Search time 27014 Seconds

(without alignments)
5433.746 Million cell updates/sec

Title: US-10-016-248-2

Perfect score: 16985

Sequence: 1 MAGAPPALALPCSLISDCC.....RSGPVGDPSTLPGSHSPKP 3104

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 segs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlh
-Q=cgnt_1/USPTO.spool/US10016248/runat_18102004_105459_11529/app_query.fasta.1.3271
-DB=GenBank -QFMT=faeap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowm62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=PCE -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10016248 @CGN 1.117721 @runat_18102004_105459_11529 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenBank.*

1: gb_db.*
2: gb_hg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_rts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15985	100.0	10136	6	AX685957 Sequence
2	15626.5	92.0	13113	9	AY210418 Homo sapi
3	14142.5	83.3	8010	6	AX685959 Sequence
4	11353.5	66.8	11580	9	AF333704 AF333704 Homo sapi

5	11322	66.7	11221	10	AY017475
6	10780	63.5	10433	6	AX374896
7	10737	63.2	10673	6	AX374891
8	10720	63.1	12525	6	AX374893
9	10591.5	62.4	13040	9	AB114605
10	10591.5	62.4	13148	9	AB114604
11	10552	62.1	10944	9	AY017307
12	10384	61.1	12486	9	AY013049
13	9659.5	56.9	10774	9	AB067481
14	9077.5	53.4	10074	6	AX537467
15	8997.5	53.0	9120	10	AK122567
16	7643.5	45.0	7323	6	AX306890
17	7643.5	45.0	8034	6	AX306891
18	7610.5	44.8	6409	6	AX306888
19	7159	42.1	5598	6	AX306886
20	6991	41.2	6145	6	AX306887
21	6978	41.1	5667	6	AX306889
22	6222.5	36.6	4360	9	AK127722
23	5172.5	30.5	3896	6	AX149475
24	5129	30.2	3905	6	AX149473
25	4826.5	28.4	3715	9	AY358174
26	4282.5	25.2	3915	6	CQ734404
27	4227.5	24.9	3289	9	AB067477
28	3950.5	23.3	5463	9	AB067471
29	3705	21.8	2334	6	CQ726904
30	3287.5	19.4	3490	9	AK126857
31	3171.5	18.7	3989	9	AK126936
32	3151.5	18.6	1866	6	CQ726903
33	2982.5	17.6	3212	6	CQ842383
34	2982.5	17.6	3518	9	AK122603
35	2813	16.6	3528	9	AK127129
36	2694	15.9	2609	6	AX477377
37	2639	15.5	2387	6	AX743686
38	2489.5	14.7	2949	6	CQ737049
39	1961	11.5	2003	6	CQ738841
40	1940.5	11.4	1615	6	AX492983
41	1868	11.0	11289	10	AF206329
42	1852	10.9	11230	6	AR435510
43	1852	10.9	11230	6	AX375163
44	1797.5	10.6	10878	6	AR435509
45	1787.5	10.6	10878	6	AX375161

ALIGNMENTS

RESULT 1
AX685957
LOCUS AX685957 10136 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 1 from Patent WO02064791.
ACCESSION AX685957
VERSION AX685957.1 GI:29371852
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Alsobrook II, J.P., Anderson, D.W., Burgess, C.E., Bolog, F.L.,
Casman, S.J., Colman, S.D., Edinger, S.R., Ellerman, K., Gerlach, V.,
Gorman, L., Grosse, W.M., Guo, X., Herrmann, J.L., Kexuda, R.,
Lepley, D.M., Li, L., Macdougall, J.R., Millet, I., Pena, C.E.,
Payman, J.A., Rastelli, L., Rieger, D.K., Shimkets, R.A., Smithson, G.,
Szytek, K.A., Stone, D.J., Tchernov, V.T., Vernet, C.A., Voss, E.Z.,
Zerhuzen, B.D., Zhong, H. and Zhong, W.
Proteins and nucleic acids encoding same
Patent: WO 02064791-A 1 22-AUG-2002;
Curagen Corporation (US)
Location/Qualifiers

FEATURES

source
1..10136
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	10136
Score:	16985.00	Matches:	3104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-016-248-2 (1-3104) x AX685957 (1-10136)

```

QY 1 MetAlaGlyAlaProProAlaLeuLeuProCysSerLeuIleSerAspCys 20
DB 1 ATGGGGGGGGCCCCCTCCCGCTTGCTGCTGCTTCAGTTGATTCAGATCTGCT 60
QY 21 AlaSerAsnGlnArgHisSerValGlyValGlyProSerGluLeuValLysGlnIle 40
DB 61 GGTACCAATACAGCGACTCCGTGGGCGTAGGACCTCCGAGCTAAGTACAGAACCAATT 120
QY 41 GluLeuLysSerArgLysValLysLeuMetProSerLysAspAsnSerGlnLysThrSer 60
DB 121 GAGTTGAAGTCTCCAGAGTGTGAGTGAAGTATGCCAGCAAGACAAACAGCCAGAGAGCTCT 180
QY 61 ValLeuThrGlnValGlyValSerGlnGlyHisAsnMetCysProAspProGlyLysPro 80
DB 181 GTGTTAACTCAGTTGGTGTGTCCCAAGGACATATATGTGTCCAGACCTGGCATATCC 240
QY 81 GluArgGlyLysArgLeuGlySerAspPheArgLeuGlySerSerValGlnPheThrCys 100
DB 241 GAAAGGGGCAAAAGACTAGGCTCGAATTTTCAGGTAGAGATCCAGGTCAGATTCACTGTC 300
QY 101 AsnGluGlyTyraPLeuGlnGlySerLysArgLleThrCysMetLysValSerAspMet 120
DB 301 AACAGGGGCTATGACTGCAAGGGGTCCAGGGGATCACTGTATGAAGAGAGCGACATG 360
QY 121 PheAlaIleTyrSerAspHisArgProValCysArgAlaArgMetCysAspAlaHisLeu 140
DB 361 TTTCGGGCTGAGACGACCAAGGCGAGTCCGAGACCCGACGATGTGTATGCCCACTT 420
QY 141 ArgGlyProSerGlyLleIleThrSerProAsnPheProIleGlnTyraPAsnAsnAla 160
DB 421 CGAGGCCCTCGGAGCATCATCACTCCCAATTTCCCACTTACAGTATGACAAACATGCA 480
QY 161 HisGlyValTyrLleIleThrAlaLeuAsnProSerLysValLleLysLeuAlaPheGln 180
DB 481 CACTGTGTGAGACATCAAGCACTCAACCCCTCCAGGGTGAATCAAGCTCGCTTGAAG 540
QY 181 GluPheAspLeuGluArgGlyTyraPThrLeuThrValGlyAspGlyGlnAspGly 200
DB 541 GAGTTGATTTGAGAGGGGCTATGACACCTGACGGTCCGATGATGATGATGATGATGATG 600
QY 201 AspGlnLysThrValLeuTyraMetSerGlnAsnAlaCysSerAspSerProHisThrPro 220
DB 601 GACCAGAGAGACAGTTCTCTCATGTCATAAATGCCGAGTGAAGCCCTCAACCCCA 660
QY 221 GlySerArgLleProGluSerMetSerGlyAspLleTyrArgGlnLysTyrThrValLeu 240
DB 661 GGGCTTCGACATCCAGAGAGCATGTCGGGACATGTGAGGAGAGAAAGAGCTGATCT 720
QY 241 GluIleCysArgAspLleSerSerSerAspAlaArgSerGlySerValArgLysSerPro 260
DB 721 GAGATCTGTGTCATGACATTGACATTCAGATGACAGTTCAGTTCAGTTCAGTTCAGTTC 780
QY 261 LysThrSerAsnAlaValGluLeuValAlaProGlyTyrGluIleGlnLysSerCys 280
DB 781 AAGACTTCAATACCTGTGAACTTGTGCTCTCCGAGAGACAGATCAGAGAGGAGGAGTTC 840
QY 281 GlyAspProGlyLleProAlaTyrGlyArgArgGlnGlySerArgPheHisArgLysP 300
DB 841 GGTGACCTCGACATCTGATGAGCCGAGAGAGGCTCCCGGTTTCCACACGATGAC 900
QY 301 ThrLeuLysPheGlnCysGlnProAlaPheGlnLeuValGlyGlnLysAlaIleThrCys 320

```

```

DB 901 ACACTCAAGTTGAGTCCAGGCCGCTTTGAGCTGTGGAGACAGAGCAATCACATGC 960
QY 321 GlnLysAsnAsnGlnTyrPserAlaLysLysProGlyCysValPheSerCysPheAsn 340
DB 961 CAABAAGATACCAATAGTGTGGCTTAAGAGCCAGGCTGGGTTCCTCGCTTCTTCAAC 1020
QY 341 PheThrSerProSerGlyValValLeuSerProAsnTyraProGluAspTyrGlyAsnHis 360
DB 1021 TTCACAGGCCGCTGGGGTGTCTGTCTCCCACTACCAAGAGACTATGGCAACAC 1080
QY 361 LeuHisCysValTyrLeuLleLeuAlaArgProGluSerArgLleHisLeuAlaPheAsn 380
DB 1081 CTCACATGTGTCTGGCTCATCTGGCCAGGCTTGAGAGCCGATCCACCTGCTTCAAC 1140
QY 381 AspLleAspValGluProGlnPheAspPheLeuValLleLysAspGlyValAlaThrAlaGln 400
DB 1141 GACATTGACGTGAGGCTCAGTTGATTTCTCGGTCACTCAAGAGAGGGGCCACCCCGAG 1200
QY 401 AlaProValLeuGlyThrPheSerGlyAsnGlnLeuProSerSerLleThrSerSerGly 420
DB 1201 GGGCCGCTCTGGGACCTTCTCAGAGAAACAGCTTCCCTCCATCAACAAGCAGTGC 1260
QY 421 HisValAlaArgLeuGluPheGlnThrAspHisSerThrGlyLysArgGlyPheAsnLle 440
DB 1261 CAGTGTGGCCGCTCGATTCAGACTCACTCACTCAAGAGAGAGGGGCTTCAACATC 1320
QY 441 ThrPheThrThrPheArgHisAsnGluCysAspProAspProGlyValProValAsnGlyLys 460
DB 1321 ACTTTTACACCTTCGACACACAGAGGCGGATCTGGCGTTCCGTAATATGGCAAA 1380
QY 461 ArgPheGlyAspSerLeuGlnLeuGlySerSerLleSerPheLeuCysAspGlyLysPhe 480
DB 1381 CGGTTTGGGAGAGCTTCACCTGAGGCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
QY 481 LeuGlyThrGlnGlySerGluThrLleThrCysValLeuLysGlnGlySerValValTyr 500
DB 1441 CTTGGAGCTCAGGGGCTCAGAACATCACTCGGCTCTCGAAGAGAGGCGAGGTGTCTGG 1500
QY 501 AsnSerAlaValLeuArgCysGluAlaProCysGlyGlyHisLeuThrSerProSerGly 520
DB 1501 AACAGGCTGTGCTGGGTGTGAAGCTCCGCTGTGTGTCTCACTTCGCTCCGACGCGC 1560
QY 521 ThrLleLeuSerProGlyTyrProGlyPheTyrLysAspAlaLeuSerCysAlaTyrVal 540
DB 1561 ACGATCCCTCTCCGGGCTGGGCTCGCTTCAAGAGATGCTTGAAGTGTGCTTGGAGT 1620
QY 541 IleGluAlaGlnProGlyTyrProLleLysLleThrPheAspArgPheLysThrGluVal 560
DB 1621 ATTGAGGCCAGCGAGGCTACCCCATCAAAATCACTTCGACAGATTCAAAACCGAGGTC 1680
QY 561 AsnTyraPThrLeuGluValArgAspGlyArgThrTyrSerAlaProLeuLleGlyVal 580
DB 1681 AACTATGACACCTCGAAGATGAGGAGGCGGAGCTTAACAGAGCCCTTGAATCGGGGTT 1740
QY 581 TyrHisGlyThrGlnValProGlnPheLeuLleSerThrSerAsnTyraLysTyraLeu 600
DB 1741 TACCACGGAGACCGAGTTCCCGATTCCTCATGACAGACAGCAACTACCTTCACTCTC 1800
QY 601 PheSerThrAspLysSerHisSerAspLleGlyPheGlnLeuArgTyrGluThrLleThr 620
DB 1801 TTCTTACGACAGAGATCACTCGGACATCGGCTTCAGCTCCGCTAATGAGACTATAC 1860
QY 621 LeuGlnSerAspHisCysLeuAspProGlyLleProValAsnGlyGlnArgHisArgLys 640
DB 1861 CTGAGATCAACACACTCTCTGATTCAGAGATCCAGTAAATGACAGGCTCATGGGAAT 1920
QY 641 AspPheTyrValGlyAlaLeuValThrPheSerCysAspSerGlyTyrThrLeuSerAsp 660
DB 1921 GACTTCAACGTGGGCGGCTGTGACCTCAAGCTGAGACTCGGGGCTACACATTAATGAC 1980
QY 661 GlyLysProLeuGlnCysGlnProAsnPheGlnTyrPheArgAlaLeuProSerCysGly 680
DB 1981 GGGGAGCTCTGAGTGTGAGCCCACTTCAGTGAAGCCGCGGCTCTGCCAAGTGTGAA 2040

```


QY 681 AlaLeuCyGlyGlyPheIleGlnGlySerSerGlyThrIleLeuSerProGlyPhePro 700
Db 2041 GCTCTCTGTGGTCTCATTTCAAGGCTCCAGTGGACCATTTGTGCCAGGGTCTCT 2100
QY 701 AspPheTyPProAsnAsnLeuAsnCysThrTriIleIleGlnThrSerIleGlyGly 720
Db 2101 GACTTCAACCCCAACAATTGAACCTGCACCTGGATTATCGAAACATCTCATGGCAAGGT 2160
QY 721 ValPhePheThrPheHleThrPheHleLeuGlnSerGlyHleAspTyPLeuLeuIleThr 740
Db 2161 GTGTCTTCACTTCCACACCTTCCACCTGGAAAGTGGCATGACTTCCCTCATCTACT 2220
QY 741 GluAsnGlySerPheThrGlnProLeuArgIleuThrGlySerArgLeuProAlaPro 760
Db 2221 GAGAACGGACGCTTACCCAGCCCTCAGGAGCTTAAGATTCCTGGCTGCCAGCTTCC 2280
QY 761 ILeSerIleGlyLeuTyPArgIleAsnPheThrAlaGlnValArgPheIleSerAspPheSer 780
Db 2281 ATCAGCGCTGGGCTCTATGGCACTTCACTGCCAGGTCCGCTTCACTCTGATTTCTCC 2340
QY 781 MetSerTyPArgIleGlyPheAsnIleThrPheSerGlyTyPArgLeuGlnProCysGlyGln 800
Db 2341 ATGTCATATGAGGATTCAACATCACCTTCTCAGAGTACGACTTGGAGCCCTGAGAGAG 2400
QY 801 ProGlnValProAlaTyPserIleArgIleGlyLeuGlnPheGlyValIleAspThrLeu 820
Db 2401 CCGAGGTCCAGCTTACAGATCCGAGAGGCTTGAGTTGGCTGGGCGACACTTG 2460
QY 821 ThrPheSerCysPheProGlyTyPArgLeuGlnGlyThrAlaArgIleThrCysLeuGly 840
Db 2461 ACCTTCTCTGCTTCCCGGGTACCGTCTGGAGGGACCCGCCCATCAGTGTCTGGG 2520
QY 841 GlyArgArgArgLeuTyPserSerProLeuProArgCysValAlaGlyCysGlyAsnSer 860
Db 2521 GGGAGAGGGGCGCTGTGAGGTCCCTCTGCGCAAGGTGTGTGTGTGTGTGTGTGTGTGT 2580
QY 861 ValThrGlyThrGlnGlyThrIleLeuLeuSerProAsnPheProValAsnTyPAsnAsn 880
Db 2581 GTCAAGGACACTCGGGTACTTGTCTGTCCCAACTTCTCTGAACTCAATACAT 2640
QY 881 HleGlnCysIleTyPserIleGlnThrGlnProGlyTyPArgIleGlnLeuValAlaArg 900
Db 2641 CATGAATGCATCTACTCATCCAGACCCAGACGAGGAAGGAATTCAGCTGAAGCCAGG 2700
QY 901 AlaPheGlnLeuSerGlnGlyAspValLeuValTyPArgIleAsnAsnSerAla 920
Db 2701 GCATTCCAACTCTCCGAAGGAGATGCTCTCAAGGTTATGATGGCAACAACCTCCGCC 2760
QY 921 ArgLeuLeuGlyValPheSerHleSerGlnMetMetGlyValThrLeuAsnSerThrSer 940
Db 2761 CGTTTGTGGAGATTGATTAAGCATTCAGATGATGGAGGACTTTGAAACAGACATCC 2820
QY 941 SerSerLeuTyPLeuAspPheIleThrAspAlaGlnAsnThrSerIleGlyPheGlnLeu 960
Db 2821 AGCAGTCTGTGGCTTGAATTCATCACTGATCTGAAAACACACGAGGCTTTGAACGTG 2880
QY 961 HlePheSerSerPheGlnLeuIleCysGlnAspProGlyThrProLysPheGlyTyP 980
Db 2881 CACTTTTCCAGCTTGAACATCATCAAAATGTGAGGACCCAGGAATCCCAAGTTTGGCTAC 2940
QY 981 LysValHleAspGlyGlyHlePheAlaGlySerSerAlaSerPheSerCysAspProGly 1000
Db 2941 AAGGTATGATGTAAGGTCATTTTGGAGGAGCTCGTCTTCACTGAGCTGACCTCGGA 3000
QY 1001 TyPserLeuArgIleSerGlnGlnLeuLeuCysLeuSerGlyGlnValArgTyPThrAsp 1020
Db 3001 TACAGCTTGGCGGTAATGAGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
QY 1021 ArgProLeuProThrCysValAlaGlyCysGlyGlyThrValAlaArgIleGlnValSerGly 1040
Db 3061 CGGCTCTGGCCCACTGT 3120

QY 1041 GlnValLeuSerProGlyTyPProAlaProTyPArgIleHleAsnLeuAsnCysIleTyPThr 1060
Db 3121 CAGGTCTGTGACCCGGGTATCCAGCTCTCTTAAAGAACATCTCAACTGCATCTGGACC 3180
QY 1061 IleGlnAlaGlnAlaGlyCysThrIleGlyLeuHlePheLeuValPheAspThrGlnGln 1080
Db 3181 ATCGAAGCAGAGGCGCGGTGCACATGGGGCTCACCTTCTGTGTGTGTGTGTGTGTGTGTGTGT 3240
QY 1081 ValHleAspValLeuArgIleTyPAspGlyProValGlnSerGlyValLeuLeuTyPArg 1100
Db 3241 GTTCAGCAGCTGTGTGCAATCTGGAGTGGGCTGTGAGAGCGGGTTCGTGTGAAGAG 3300
QY 1101 LeuSerGlyProAlaLeuProLysAspLeuHleSerThrPheAsnSerValLeuGln 1120
Db 3301 CTGAGTGGCGGCGCTTGCACCAAGACCTGCATAGCACTTCAACTCGGTCTGTGGAG 3360
QY 1121 PheSerThrAspPheThrSerTyPArgIlePheAlaIleGlnPheSerValSerThr 1140
Db 3361 TTCAGCAGTGACTTCTTCAACAGCAAGCAGGCGCTTGCATTCATTTCAAGTGTCCACA 3420
QY 1141 AlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerTyPArgSerGlyAspSerThr 1160
Db 3421 GCACGCTCTGCAATGACCTTGGAAATCCGCAAAATGGAGTGGAGTGTGACAGTTGG 3480
QY 1161 GluAlaGlyAspSerThrValPheGlnCysAspProGlyTyPArgIleGlnGlnIleSerAla 1180
Db 3481 GAAGCGGCGACCTCCACAGTGTTCAGTGTGACCTTGTGCTGACGCGCTCAGAGGAAGTGA 3540
QY 1181 ILeSerCysCysValIleGlnAsnArgPhePheThrGlnProSerProProThrCys 1200
Db 3541 GAGATCAGCTGTGTGAAGATCGAAGACAGGTCTTCTGTGGAGCGCCAGCCGCAACATGC 3600
QY 1201 IleAlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeuSerProAsnTyP 1220
Db 3601 ATGCTCTCTGCGGGAGACCTGCAGACGACATCTGAGATCATCTTCAACCAATATAC 3660
QY 1221 ProGlnProTyPProProGlyTyPArgIleGlyCysAspTyPLeValThrValSerProAspTyP 1240
Db 3661 CCAGAACCTTACCCCGCAGGACGAGAGTGTGATGGAAATGACCGTCTACCAAGACTAC 3720
QY 1241 ValIleAlaLeuValPheAsnIlePheAsnLeuGlnProGlyTyPArgPheLeuHleIle 1260
Db 3721 GTCATGCGCTGTATTTAACTTTAACTTTAACTTGAACCTGAGCCTGTGATGACTTCCATATC 3780
QY 1261 TyPArgGlyArgAspSerLeuSerProLeuIleGlySerPheTyPArgIleGlnPro 1280
Db 3781 TACGAGGAGAGGAGACTTCTCAGCCCTCATAGAGCTTCTATGGCTCCAGCTCCA 3840
QY 1281 GlyArgIleGlySerSerSerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerVal 1300
Db 3841 GGCCGATTTGAAGCAGACGACCAACAGCTTCTTCTGCGCTTCCGACGATGATCTGTG 3900
QY 1301 SerAsnAlaGlyPheValIleAspTyPThrGlnAsnProArgIleuSerCysPheAspPro 1320
Db 3901 AGCAATGCTGTGCTTCTGTCATTTGACTATACAGAAACCCGGGAGATCATGTTTGAATCT 3960
QY 1321 GlySerIleLysAsnGlyThrArgValGlySerAspLeuLysLeuGlySerSerValThr 1340
Db 3961 GGTTCATCAAGAACGAGCACACGGGTGGGGTCCGACTGAAAGTGGGCTCTCCGTACCC 4020
QY 1341 TyPTyPAsnIleGlyGlyTyPArgIleValGlnGlyThrSerThrLeuSerCysIleLeuGly 1360
Db 4021 TACTTACTGCCAGGGGCTTACAGATTTGAGGAGCACTTCGACCTTGAGCTGCACTTGGGG 4080
QY 1361 ProAspGlyLysProValTyPAsnAsnProArgProValCysThrAlaProCysGlyGly 1380
Db 4081 CTGATGGGAAAGCCCGGTGTGAACAATCCCGGCGAGTGTGCACAGCCCTGTGGGGGA 4140
QY 1381 GlnTyPValGlySerAspGlyValValLeuLeuSerProAsnTyPProGlnAsnTyPThrSer 1400
Db 4141 CAGTATGTGGGTTCGAGCAGGAGT 4200
QY 1401 GlyGlnIleCysLeuTyPThrValThrValProLysAspTyPValPheGlyGlnPhe 1420

Db	4201	GACAGATCGCTGTAATTTTGTACTGTGCGCCAGAGACTATGTGTGTTGGCCAGTTC	4260
Qy	1421	AlaPhePheHisThrAlaLeuAspValValGluValHisAspGlyHisSerSerGlnHis	1440
Db	4261	GCCTTTTTCACACGCGCCCTCAACGACGTGTGAGGTTCAACGAGGCGCACGACGAC	4320
Qy	1441	SerArgLeuLeuSerSerLeuSerGlySerHisSerHisGlyLeuSerLeuProLeuAlaThr	1460
Db	4321	TGCGGCTCTCTACGCTCCCTCTCGGGCTCCCAACAGAGAACTACCTCCCTTGGCCACC	4380
Qy	1461	SerAsnGlnValLeuLeuLeuPheSerSerAlaGlyGlyLeuAlaProAlaArgGlyPheHis	1480
Db	4381	TCCAACTCAAGTCTCATTAAGTTCAAGCGCCAAAGGCTTCGACCGACCGAGGCTTCCAC	4440
Qy	1481	PheValTyrGlnAlaValProArgHisSerAlaThrGlnCysSerSerValProGluPro	1500
Db	4441	TTTGTCTACCAAGCGGTTCTTGAAACAGCGCCAGCGACGTGAGTCTGTGCTCGGAAACC	4500
Qy	1501	ArgTyrGlyLysArgLeuGlySerAspPheSerValGlyValAlaIleValArgPheGluCys	1520
Db	4501	CGCTATGGCAAGAGGCTGGGAGTGACTTCTCGGTGGGGGCAATGTCGCTTGAAATGC	4560
Qy	1521	AsnSerGlyTyrAlaLeuGlnGlySerProGluIleGlnCysLeuProValProGlyVala	1540
Db	4561	AACTCGGCTATGCGCTCGAGGGGTGCGCAAGATCGAGTGCCTCCCTGTGCTGGGGCC	4620
Qy	1541	LeuAlaGlnTyrPheAsnValSerAlaProThrCysValValProCysGlyGlyAsnLeuThr	1560
Db	4621	TTGGCGCCAAATGGAATGTTCTACGCGCCCAAGTGTGGTGGCGTGTGAAGCAACCTCACA	4680
Qy	1561	GluArgArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeuAsn	1580
Db	4681	GAGCGCAAGGGGACACCATCTGTCCCTGTGGCTTCCAGAGCGGTACCTCAACAGCTTAC	4740
Qy	1581	CysValTyrPheValIleValProGluGlyAlaGlyIleGlnIleGlnValValSerPhe	1600
Db	4741	TGTGTGTGAAGATGTGTGTCCGAAAGCGCTGGATCCAGATCCAAAGTTGTCAAGTTT	4800
Qy	1601	ValThrGluGlnAsnTyrPheSerLeuGluValPheAspGlyAlaAspAsnThrValThr	1620
Db	4801	GTGACAGAGCAGACTGGGACTCGCTGAGATTTTATGTGTCAATTAACCTGTAAACC	4860
Qy	1621	MetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerThrSerAsnGln	1640
Db	4861	ATGTGTGAGATTTCTCAGAAACAAACCGTCTGCTCTTGAAACAGCACTCCAAACGAG	4920
Qy	1641	LeuTyrLeuHisPheTyrSerAspIleSerValSerAlaIleGlyPheHisIleGlnIleTyr	1660
Db	4921	CTCACTTCTCATTTCTCATCTCAGATATCAGCTATCTGCAAGCTGCTTCCACTTGGAGTAC	4980
Qy	1661	LysThrValGlyLeuSerSerCysProGluProAlaValProSerAsnGlyValLysThr	1680
Db	4981	AAAAACGCTGGGCTGAGCACTTGTCCGGAACCTGCTGTCAGCAAGGGGTGAAGACT	5040
Qy	1681	GlyGluArgTyrLeuValAsnAspValValSerPheGlnCysGluProGlyTyrAlaLeu	1700
Db	5041	GCGGAGGCGTACTTGTGTGAATGTGTGTCTTTCACGTGAGCGGGAATATGCCCTTC	5100
Qy	1701	GlnGlyHisAlaHisIleSerCysMetProGlyThrValArgArgTyrAsnTyrProPro	1720
Db	5101	CAGGGCCACGCCCACTCTCTGCAATGCCGGAACAGTCCGGCGATGAACTACCTCTCT	5160
Qy	1721	ProLeuCysIleAlaGlnCysGlyGlyThrValGluGluMetGluGlyValIleLeuSer	1740
Db	5161	CCACTCTGATTGCAAGTGTGGGGAACAGTGTGAGAGATGAGAGGGGTGATCTCTGAGC	5220
Qy	1741	ProGlyPheProGlyAsnTyrProSerAsnMetAspCysSerTyrLysIleAlaLeuPro	1760
Db	5221	CCGGGCTTCCAGGCACTAACCCCAATACATGAGTGTCTCTGGAATAATGCACTGGCC	5280
Qy	1761	ValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGluProAsnHisAspTyr	1780
Db	5281	GTGGGCTTTGGAGCTCACTACAGTTCCTGAATCTTCCACCGAGCCCAACGACGACTAC	5340
Qy	1781	IleGluIleIleArgAsnGlyProTyrGluThrSerArgMetMetGlyArgPheSerGlySer	1800
Db	5341	ATGGAATTCGGATGGATGCCCTTATGTAGACACCGCATGATGTGAAGATTCACTGTGAAAC	5400
Qy	1801	GluLeuProSerSerLeuLeuSerThrSerHisGlyThrThrValTyrPheHisSerAsp	1820
Db	5401	GAGCTTCCAAAGCTCCCTCTCTCCAGTCCAGGACCGACCAACCGTGTATTTCCACAGCAC	5460
Qy	1821	HisSerGlnAsnArgProGlyPheLysLeuGlyTyrGlnAlaIleTyrGluLeuGlnIleCys	1840
Db	5461	CACTCCCAAAATCGGCAAGATTCAGAGTGTGAGTATCAGGCTTATGAATTCAGAGTGC	5520
Qy	1841	ProAspProGluProPheAlaAsnGlyIleValArgGlyValGlyTyrAsnValGlyGln	1860
Db	5521	CCAGACCCAGAGCCCTTTCCTCCATGTGATGTGAGGGAGCTGGCTTACAACTGGGACAA	5580
Qy	1861	SerValThrPheGluCysLeuProGlyTyrGlnLeuThrGlyHisAspValLeuThrCys	1880
Db	5581	TCAATGACCTTGAATGTGCTCCCGGGGTATCAATGACTGGCAACCTGTCTCACTGT	5640
Qy	1881	GlnHisGlyThrAsnArgAsnThrAspHisAspProLeuProLysCysGluValProCysGly	1900
Db	5641	CAACATGGCAACCAAGGAGCTGGAGCAACCCCTGCGCAAGTGTGAAGTCTTGTGGC	5700
Qy	1901	GlyAsnIleThrSerSerAsnGlyThrValTyrSerProGlyPheProSerProTyrSer	1920
Db	5701	GGGAACATCACTTCTTCCAGAGGAGCTGTATCTCCCGGGGTTCCCTAGCCGTACTCC	5760
Qy	1921	SerSerGlnAspCysValTyrLeuIleThrValProIleGlyHisGlyValArgLeuAsn	1940
Db	5761	ACCTCCAGAGCTGTGTGTGCTGATCACTGTGCCATTTGGCCATGGCGTCCGCTCAAC	5820
Qy	1941	LeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTyrAspGlyProGln	1960
Db	5821	CTCAGCTGTGAGAGACAGAGCCCTGTGAGATTTCACTCACTGTGAGATGGGCAACAG	5880
Qy	1961	GlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaLysLysThrValGlnSer	1980
Db	5881	CAACAGCACACCGCTCGGCTTCAACCGGAGCATGGCCAAAGAAACAGTGCACAGT	5940
Qy	1981	SerSerAsnGlnValIleLeuLysPheHisArgAspAlaIleThrGlyGlyIlePheAla	2000
Db	5941	TCACTCAACAGAGCTCTGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCC	6000
Qy	2001	IleAlaPheSerAlaTyrProLeuThrLysCysProProProThrIleLeuProAsnAla	2020
Db	6001	ATAGCTTTCCTCGCTTATCCACTCAACCAATGCTCTCTCCACATCTCTCCCAAGCC	6060
Qy	2021	GluValValThrGluAsnGluGluPheAsnIleGlyAspIleValArgTyrArgCysLeu	2040
Db	6061	GAATCGTCAAGAGATGGAATTCATATATGTGACATCTTACGCTTACAGATGCTTC	6120
Qy	2041	ProGlyPheThrLeuValGlyAsnGluIleLeuThrCysLysLeuGlnIleThrGln	2060
Db	6121	CTGTGCTTTCCTTATGTGGGAGATGAAATTTGACTGACCTTGAAACTTGAACCTTACG	6180
Qy	2061	PheGluGlyProProProIleCysGluValHisCysProThrAsnGluLeuLeuThrAsp	2080
Db	6181	TTTGAAAGAACACCCCGCATATGTGAAGTGCATCTGCCAAATAGAGCTTGTGACAGAC	6240
Qy	2081	SerThrGlyValIleLeuSerGlnSerTyrProGlySerTyrProGlnPheGlnThrCys	2100
Db	6241	TCCACAGGCGTATCTCTGAGCCAGAGCTTACCTCGAAGATATCCCAAGTTCCAGACCTGC	6300
Qy	2101	SerTyrLeuValArgValGluProAspTyrAsnIleSerLeuThrValGluTyrPheLeu	2120
Db	6301	TCTTGGCTGTGAGATGTGAGCCCGCATTAATCATCTCTCTCACTGAGATCTTCTCTC	6360
Qy	2121	SerGluLysGlnTyrAspGluPheGluIlePheAspGlyProSerGlyGlnSerProLeu	2140
Db	6361	AGCGAAGAACAAATATGATGATTTTGAATTTTGTGATGTGCTCATCAGACAGAGTCTCTG	6420

Qy	2141	LeuValAlaLeuSerGlyAsnThrSerAlaProLeuIleValThrSerSerAsnSer	2160	Qy	2501	LeuThrGlnGlyAsnGlnPheAsnLeuAsnAspValValIysPheValCysAsnProGly	2520
Db	6421	CTGAAGCCCTCACTAGGAAATTACTCAAGCTCCCTGATTGTACACAGCTCAAGCACTCT	6480	Db	7501	CTCACTAGGGTAACAGTTTAACCTCAAGATGTGTTCAAGTTTGTTCGAACCTCGG	7560
Qy	2161	ValThrLeuArgThrPheSerAspHisAlaThrAsnArgLysGlyPheLysIleArgThr	2180	Qy	2521	TyrMetAlaGlnGlyAlaAlaArgSerGlnCysLeuAlaSerGlyGlnTrpSerAspMet	2540
Db	6481	GTCGACCTGGCTGTGATCTGATCAGCCCTACATGGGAAGGGCTTCAGATCCGCTAT	6540	Db	7561	TATATGCTAGGGGGCTGCTAGGTCCCAATGCTGGCCAGGGGCAATGGAGTGCATG	7620
Qy	2181	SerAlaProThrCysSerLeuProArgAlaProLeuHisGlyPheIleLeuGlyGlnThr	2200	Qy	2541	LeuProThrCysArgGlnIleAsnCysThrAspProGlyHisGlnGlnAsnSerValArg	2560
Db	6541	TCAGCCCTTACTCAGCTGCGCCAGGGCTCCACTCCATGGCTTCATCTAGGCAAGACC	6600	Db	7621	CTGCCCACTGGCAAAATCAATCACTGACAGATCTTGGAACCAAGAAATAGATTGCT	7680
Qy	2201	SerThrGlnProGlyGlySerIleHisPheGlyCysAsnAlaGlyTyrArgLeuValGly	2220	Qy	2561	GlnValHisAlaSerGlyProHisArgPheSerPheGlyThrThrValSerTyrArgCys	2580
Db	6601	AGCACCAAGCCCGGGGCTCCATCCATTGTGGCAAGCGGGCTACCGCTGTGGTGGGA	6660	Db	7681	CAGGTCCAGCCGACGGGCGCCGACAGGTTCAGCTTGGCAACCACTGTCTTACCGGTCC	7740
Qy	2221	HisSerMetAlaIleCysThrArgHisProGlnGlyTyrHisLeuTrpSerGluAlaIle	2240	Qy	2581	AsnHisGlyPheTyrLeuLeuGlyThrProValLeuSerCysGlnGlyAspGlyThrTrp	2600
Db	6661	CACAGCATGGCCATCTGTACCCGGCACCCCAGGGCTACCACTGTGGAGGGAAGCCATC	6720	Db	7741	AACCAAGCTTTTCACTCTGGGCACCCAAGTGTCTAGCTGACGAGGAGATGGCACAATGG	7800
Qy	2241	ProLeuCysGlnAlaLeuSerCysGlyLeuProGluAlaProIysAsnGlyMetValPhe	2260	Qy	2601	AspArgProArgProGlnCysLeuLeuValSerCysGlyHisProGlySerProProHis	2620
Db	6721	CTCTCTGTCAAGCTCTTCTCTGTGGCTTCTGAGGGCCCCAAGAAATGGAATGTGT	6780	Db	7801	GACCGTCCCGCCCGCCAGTGTCTTGTGTCTGTGTCCATCCGGGCTCCCGCTCAC	7860
Qy	2261	GlyLeuGlnThrThrValGlyThrThrAlaValTyrSerCysSerGlyGlyTyrHisLeu	2280	Qy	2621	SerGlnMetSerGlyAspSerTyrThrValGlyValAlaValArgTyrSerCysIleGly	2640
Db	6781	GGCAGAGGATACACAGTGGGAACCAAGCCCTGTACAGCTGCAATGAAGGTACCACTC	6840	Db	7861	TCCCAAGTGTCTGGAGCAATTATCTGTGGAGCAATGTGGGTGCACTGATGGC	7920
Qy	2281	GlnAlaGlyAlaGlnAlaThrAlaGlyCysLeuAspThrGlyLeuTrpSerAsnArgAsn	2300	Qy	2641	LysArgThrLeuValGlyAsnSerThrArgMetCysGlyLeuAspGlyHisTrpThrGly	2660
Db	6841	CAGCAGGGCGGTGAGGCCATCGACAGTGTGACACAGGCCCTATGGAGCAACCGCAT	6900	Db	7921	AAGGTACTCTGGGGGAACAGCACCCGATGTGGGTGGATGACACTGACATGGC	7980
Qy	2301	ValProProGlnCysValProValThrCysProAspValSerSerIleSerValGluHis	2320	Qy	2661	SerLeuProHisCysSerGlyThrThrSerValGlyValCysGlyAspProGlyIleProAla	2680
Db	6901	GTCCACACACAGTGTCTCTGTGATCTGTCTGATGTCAATACAGAGTGTGACAT	6960	Db	7981	TCCCTCCCTCACTGTAGAACCAAGGTGGAGATTGTCCGTGACCTCGGATCCCGCT	8040
Qy	2321	GlyArgThrPheArgLeuIlePheGlnThrGlnThrGlnPheGlnAlaGlnMetLeuIle	2340	Qy	2681	HisGlyIleArgLeuGlyAspSerPheAspProGlyThrValMetArgPheSerCysGlu	2700
Db	6961	GGCGGATGGAGGCTTATCTTTGAGACACAGTATCAAGGCCCAAGCTGATCTCATC	7020	Db	8041	CATGGCATCCGTTTGGGGGACAGCTTGTATTCAGGCACTGTGATGCCCTTCACTGTGAA	8100
Qy	2341	CysAspProGlyTyrTyrTyrThrGlyGlnArgValIleArgCysGlnAlaAsnGlyLys	2360	Qy	2701	AlaGlyHisValLeuArgGlySerSerGlyAspGlnThrCysGlnAlaAsnGlySerTrpSer	2720
Db	7021	TGTACCTCTGGCTACTACTATACCTGGCCAAAGGGTCACTCCCTGTCAAGGCCAATG	7080	Db	8101	GCTGGCAGCGTCCGGGGATGCTGACAGCGGCACTGTCAAGCCAAATGGCTGTGAGC	8160
Qy	2361	TrpSerLeuGlyAspSerThrProThrCysArgIleIleSerCysGlyGlyLeuProIle	2380	Qy	2721	GlySerGlnProGlyCysGlyValIleSerCysGlyAsnProGlyThrProSerAsnAla	2740
Db	7081	TGGAGCTCGGGGACTACGCCCACTGCCAATCATCTCTGTGAGAGCTCCGAT	7140	Db	8161	GGCTCGCAGGCTGATGTGAGTGTCTTGTGGGAACCTGGGACTCCAAAGTAAAGCC	8220
Qy	2381	ProProAsnGlyHisArgIleGlyThrLeuSerValTyrGlyAlaThrAlaIlePheSer	2400	Qy	2741	ArgValValPheSerAspGlyLeuValPheSerSerIleValTyrGlnCysArgGlu	2760
Db	7141	CCCCCAATGGCCACCGCATCGGAACACTGTGTCTTACGGGGCAACCGCATTTCTCC	7200	Db	8221	CGAGTTGTTCAGTATGGCTGTGTTTCTCCAGCTCTATCGTATGAGTGCAGGGA	8280
Qy	2401	CysAsnSerGlyTyrThrLeuValGlySerArgValArgGlyCysMetAlaAsnGlyLeu	2420	Qy	2761	GlyTyrTyrAlaThrGlyLeuLeuSerArgHisCysSerValAsnGlyThrTrpThrGly	2780
Db	7201	TGCATTCGCGATACACACTGTGTGGCTCCAGGGTGTGATGTCAATGGCCAAATGG	7260	Db	8281	GGAATCAACCCACAGGCTGTGACGCGTCACTGCTGGTCAATGTATCTGGACAGGC	8340
Qy	2421	TrpSerGlySerGlyValArgCysLeuAlaGlyHisCysGlyThrProGluProIleVal	2440	Qy	2781	SerAspProGlyCysLeuValIleAsnCysGlyAspProGlyIleProAlaAsnGlyLeu	2800
Db	7261	TGGAGTGGCTCTGAAGTCCGCTGTGCTGTGACACTGTGGGACTCCAGAACCATGTCTC	7320	Db	8341	AGTGAACCTGAGTCCCTGTGATTAACCTGTGTGACCTGTGGATTCAGGCAATGGCTT	8400
Qy	2441	AsnGlyHisIleAsnGlyGlyAsnThrSerTyrArgGlySerValValTyrGlnCysAsn	2460	Qy	2801	ArgLeuGlyAsnAspPheArgTyrAsnThrValThrTyrGlnCysValProGlyTyr	2820
Db	7321	AACCGACATCAATGGGGAACTACAGCTACCGGGCAGTGTGTACCAATGCAAT	7380	Db	8401	CGGCTGGCAATGACTTCAAGTACAAACAACTGTGACATATCATGTGTCTGGCTAT	8460
Qy	2461	AlaGlyPheArgLeuIleGlyMetSerValArgIleCysGlnGlnAspHisIleTrpSer	2480	Qy	2821	MetMetGlySerHisArgValSerValLeuSerCysThrLysAspArgThrTrpAsnGly	2840
Db	7381	GCTGGCTTCCGCTCGATCGGATGTCTGTGGCATCTGCCAGCAGATCATCATGTGCTG	7440	Db	8461	ATGATGAGTCAATAGATGATCTGTGTAGCTGACCAAGGACCGGACATGGAATGGA	8520
Qy	2481	GlyLysThrProPheCysValProIleThrCysGlyHisProGlyAsnProValAsnGly	2500	Qy	2841	ThrLysProValCysLysValAlaMetCysLysProProProLeuIleProAsnGlyLys	2860
Db	7441	GGCAAGACCCCTTCTGTGTGCAATTACTGTGAGACCCAGGCAACCCCTGTCAAGGCG	7500	Db	8521	ACCAAGCCCGTGTGCAAGCTCATGTGTGAAGCAACCTCGCTCAATCCCAATGGGAG	8580
Qy	2501	ValValGlySerAspPheMetTrpGlySerSerValThrTyrAlaCysLeuGlnGlyTyr	2520	Qy	2861	ValValGlySerAspPheMetTrpGlySerSerValThrTyrAlaCysLeuGlnGlyTyr	2880

Db	FEATURES
8581	
8581	GTGGTGGGGCTGACTTCACTGCGGGCTCAAGTGTGACTTATGCTGCTGGAGGGGTAC 8640
Qy	2881 GlnLeuSerLeuProAlaValPheThrCysGluGlyAsnGlySerTyrThrGlyLeu 2900
Db	8641 CAGCTCTCCCTCCCGGGGTGTTCACTGTGTGAGGAAATGGGTCTTGGACCGGAGAGCTG 8700
Qy	2901 ProGlnCysPheProValPheCysGlyAspProGlyValProSerArgGlyArgArgGlu 2920
Db	8701 CCTCAGTGTTCCTCTGTGTTCCTGGGGATCTGTGTCTCCCTCCGAGGAGAGAGAG 8760
Qy	2921 AppArgGlyPheSerTyrArgSerSerValSerPheSerCysHisProProLeuValLeu 2940
Db	8761 GACCAGGCTTCTCTACAGGCTCATCTGTCTCTCTCTCTGTCATCCCTCTGTGTGTG 8820
Qy	2941 ValGlySerProArgArgPheCysGlnSerAspGlyTyrThrPseGlyTyrGlnProSer 2960
Db	8821 GTGGGCTTCCACGAGGTTTGGCAGTCAGATGGACATGAGTGGACACCCAGCCAGC 8880
Qy	2961 CysIleAspProThrLeuThrCysAlaAspProGlyValProGlnPheGlyIleGln 2980
Db	8881 TGCAATAGTCGACCTTACACCGTGTGGGACCTGTGTGTGCCAGTTTGGATTCAG 8940
Qy	2981 AsnAsnSerGlnGlyTyrGlnValGlySerThrValLeuPheArgCysGlnGlyTyr 3000
Db	8941 AACAAATTCAGGGCTACAGAGTGGAGACAGCTCTCTCTGTCATCAAAAGGCTAC 9000
Qy	3001 LeuLeuGlnGlySerThrThrArgThrCysLeuProAsnLeuThrTyrPseGlyTyrPro 3020
Db	9001 CTGCTTACGGGCTCCACACAGACCTGCTCTCCAAACCTGACCTTGAAGTGGAAACCCCA 9060
Qy	3021 ProAspCysValProHisCysArgGlnProGlyLeuThrProThrHisAlaAsnValGly 3040
Db	9061 CCTACTGTGTCTCCCACTGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 9120
Qy	3041 AlaLeuAspLeuProSerMetGlyTyrThrLeuIleThrProAlaArgAlaAspPro 3060
Db	9121 GCCCTGTGTTTCCCTTCATGAGGCTACACGCTCATTACTCTCTGACAGAGGCTTCTCCC 9180
Qy	3061 SerArgValAlaProSerThrAlaProAlaArgArgMetAlaAlaGlyGlnAlaSerArg 3080
Db	9181 TCAGGGGTGGCTCGACACCGACCTGCAAGGGAGTGCACACTGGACAGGAGCCGC 9240
Qy	3081 ProSerAlaTyrPheSerGlyProValGlyAspProSerThrLeuProGlySerHisArg 3100
Db	9241 CCATCTCTCTGAGAGTCCGGCCCACTGTGAGACCCATCAACACTGCCCGGAGACCCGC 9300
Qy	3101 SerProLeuPro 3104
Db	9301 TCACCCAGGCT 9312
RESULT 2	
AY210418	13113 bp mRNA linear PRI 31-JUL-2003
LOCUS	Homo sapiens CUB and sushi multiple domains 2 (CSMD2) mRNA,
DEFINITION	complete cds.
ACCESSION	AY210418
VERSION	AY210418.1 GI:30908442
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Lau, W.L. and Scholnick, S.B.
JOURNAL	Identification of two new members of the CSMD gene family
REFERENCE	Genomics 82 (3), 412-415 (2003)
AUTHORS	2 (bases 1 to 13113)
JOURNAL	Lau, W.L. and Scholnick, S.B.
TITLE	Direct Submission
DATE	Submitted (02-JAN-2003) Otolaryngology - Head & Neck Surgery,
INSTITUTION	Washington University School of Medicine, 660 South Euclid Ave.,
LOCATION	Saint Louis, MO 63110, USA

FEATURES

Source

Location/Qualifiers

1..13113

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

1..13113

/gene="CSMD2"

30..10493

/gene="CSMD2"

/note="contains 14 CUB domains and 26 sushi domains"

/product="CUB and sushi multiple domains 2"

/protein_id="AA034701.1"

/db_xref="GI:30908443"

/translation="MRLISIAPEPMQVQSHHPRSAQNCSTFQLHGNCTVESPGPY
GYPNANCTWTITAEORHIOVPOSPALIEEDFVLISVDEPOPEBNIRTLTGOLP
ATVSAATLSRLISDYAVSQGHATYELPSHTCNPRGLPBGIOQSGTSPNGDK
VRSNCGLPFLGHAVLTCAGSESAITMDPLPSCRADACGTLRQSGITSSPH
PSEYNNADCTTTLAEGDTLALVFTFOLEBDYDFLEVTGESSLMFTGASLPAP
VTSKWLHLHTSDGNHRQGFSAQYQVKQIELKSGVKLMPKMDOKSVTLTV
GVSQHNMCPPDEIGPERGRULSDPLRSGVQFTCNBEGYDQSGKRTCKMSDFAA
WSDHRYCARAKMCDHLRPGSGLITSPMPPIQYNNACWMIITLANSKTIKAEF
FLERGYDTLTVGDGODDQKTVLITLTPVPLIVSTHOMWLFOYQSSSLG
FASVEIEQSGCGDPGI PAYGRREGSPHHGDTLKEPCQAPFELVCGATTCQNM
MSAKKPGVSCFFPPTSPGVLSYRNPEDYGNHLHCWMLILAPESRHLAEVDID
VERQDPLVTKDAIAEAVALGTSNQLPSSITSSGHVARELQTHSTGRFNIT
PTFRHNCPPDGIVNKRFRDLSLOLSSISFLCDEGFLTQSGETITCYLKESV
MNSAVLRCEAPCGHLLTSPGTLISPMWPGFKYKALSCAWYLEAPGPKITPFRFK
TEVNDTLEVDGRTYSAPLIGVHTGVPQPELITSTSYVLITFSTDHSDIGLOP
YETITLQSDHCDPGLPVNGORHGDPEVALVLPSCDSGYTSLDGEPLREBNQMS
RALPSCALCGGFTQSSGTLISGFPDPYNNINCTWITITSHKGVFTFTHTLE
GHDYLLITNGSTFQPLRQUTGSRILPISAGLYGNFTAOVRITSDMSYEGENIT
PSEYDLPESEPEVPAISIRKGLQPRGVDTLTFCSPYRLGARTICLGRRLMS
SPLPRCAECGNSVLTQGTLLSNPFPVNNNNHECTISLOTQPKQLKARAEIS
EGDYLKVDGNNNSARLLGVFSHEMMGVLTNLSSTSSLMDFITDAETVSGLFHS
SFEILKCEDPGLPKFGYVHDEGHAGSVSFCDPQSGYSLGLCSERLTLGSRFTWR
PLPTVACGEGTVREGVGVLSPGYVAPYNNINCTWITITSHKGVFTFTHTLE
EVDYVLRIMDPVPSGVILKELSPALPKDHLSTFNSVLTDFSTDFLSKGRALORS
VSTATSCDPLGIPONGSRSDSWAGDSYVQCPGVALQSSASISCVKLENRFPWP
SPTTCLAPCGDGLPONGSVLISPNYPERPGKCDKWTVPVLYVALVNIPLNP
GYDLPLHYDGLDLSPLIGSFYQGLRLESSNSFLAFRSDASVNAAGFVDYTE
NPRSCDPGSLIKNGTRVSGDLKLGSLVYCHQGEVGEVSTLSLGPGRVWVN
PRPVCAAPCGQVYDGVLSYSPYPNQNTSYLITFVTPKDYVVRGQAPFHTL
NDVVEYHDSGHSLSLSSGHTGSLPLATNCVLTIRPSAGLAPAKGFHTVLA
VPTSAQDCSVPRPYRGRKGLSPSGALITREPCNGVALQSGPELCELPVAFALQ
MNVAPRTCVPCGGLTERRGTLISPEPPEYLNLSNWKIVLVEBAGIQLQVSVY
TEQWMSLVEVDGADNTVLMGSPSGTVALNLSNWKIVLVEBAGIQLQVSVY
YKTVGLSCCEPAVPSNGVYTGERYLVNDVVSFCGEBVALQGHAIISCMPTVRWN
YPPPLCAQCGTVEEMGVLTSPGPNYPSMNDGSKXIALPVGPAHIOPLNSTR
PNHDYIETRNQPYETSHMRGFSGSELPSLSTSHETTYFPHSDONRPGFLBYO
DLATYSHOISSFLRGEFDLSLEPNTSPPVAAASYWMDPCCGAYELOECDFPEFANG
LYRGAGNNQGSVTFECLPGYQLTGHEVLTQGHNTNNMDHPKCEVPCGGNTSSN
GTVSPGPEPSSYSDCWLITVPLIGHGRVLMLSLQTPSGPFIITWGPQDAR
LGVTFRMARKTVQSSNOVLKFRHRAAGGIFALFASVYPLKCPPTILPNAEV
TEMBEFGITDVIYRCLPGFTLVGNELITLCKLQYFEGEPGPRCEVHCTNGLDS
TGVLSQSPGYSYPOFOTCSMLVREVDNISLTVETFESEKXODEEITFESGCGP
LKLALSGNSAPLIVTSSNSVYLRMSDHAHYRKIRYSAPYSGSLPAPLHGTLL
GOSTORGBIHFCCNAGRVLVGHSMAICTRHQGHLMSEALPLCOALSCTGLPEAK
NGMFRKELYTVTKAMYSCEGYHLQAGATVACLDITGLMSNRNVPOCVPTCDY
SSISVEHGRRLIFETYOFOQOALMICPENGSYVQYGVRCQANGMSLSDPTCR
IISGELPIPPNHRIGTLISVGTATLIPGNSGYVQYGVRCQANGMSLSDPTCR
AGHCGRPEPLVNGINGENYSYRGSVYQCNAAFRLIGMSVRIICODHHTSGTPTCV
LVSCGHPSPHSGMSDSTVGAIVVYSCIGRFLVNGSTRVTCGIDHHTSGTPTCV
GTSVGVGDBGIPAHGIRLQDSDPPTWAFRSEBAHVLRGSEERTCOAGNSGSSP
ECGVISGNGPTSNARVYFSDLVSSISVYECRGYVATGTLSSHSCSTGWTGSD
PECLVINGDPIIPANGRLGDNFRNKVTVYVCEGRYVATGTLSSHSCSTGWTGSD
TKPVCCALMCKPPLIPNGKVSDFMWSVSVYACLEGVOLSLPAVFTCEGSGWTG
ELPOCFPVEFGDPSRGRREDRGSYSVSFSGHPLVLSGRRRCQSGWTGSG
TQSCIDPLTTCADGPVPOFGIQQNSQVYSGTULPFCOCGYLLQSGTTPRCLNLT
TWGCTPDPCVPHCRQPEPTANNVGLALBSMGITLITSCOBGSLKSGSHRTCKA
DGSWTKRPICLEVRNGPINTARPPITQALIPEDVAKNSLMGALEYQKQKPA
MLRWYQVANSKVAATWIDHSGVELHAGTYKKEDPHLLLOYQITGPVELFMNFKF
DHWALDGVSSSESGATPFIYQGSVKGDFQGFQFORLRLRLLESPPESIGRFAANS

SSVAAALVPEFIALIAGFVLVLYKHRRRPKVPFNGYAGHENTVRATFNPMDYRNI
QPTDIMASEALEFTSTVCTAV"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	13113
Score:	15626.50	Matches:	2900
Percent Similarity:	92.64%	Conservative:	19
Best Local Similarity:	92.03%	Mismatches:	35
Query Match:	92.00%	Indels:	198
DB:	9	Gaps:	10

US-10-016-248-2 (1-3104) x AY210418 (1-13113)

```

Qy      2 AlaGlyAlaProProAlaLeuLeuLeuProCySerLeuIle---SerAspCys--- 19
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      826 TCTGGTTCACCGGACCGACCTCCAGCCCGCTTATACAGCAGAAAGCTGGTCGAC 885
Qy      20 CysAlaSerAengInaRghIseSerValGlyValGly-----ProSerGluLeuVal 36
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      886 TGCATTCACATCGGATGGCAACACCGGCGCGGATTCAGTCCCAATACCA--GTC 944
Qy      37 LysLysGlnIleGluLeuLysSerArgGlyValLysLeuMetProSerLysAspAsnSer 56
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      945 AAGAAAGCAATTGAGTTGAAGTCTCGAGGTGTAAGCTGATGCCAGAAAGACACAGC 1004
Qy      57 GlnLysThrSerValLeuThrGlnValGlyValSerGlnGlyHisAsnMetCysProAsp 76
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1005 CAGAAAGAGCTGTGTTAACTCAGGTGTGTGTCCAAAGACATTAATGTGTCCAGAC 1064
Qy      77 ProGlyIleProGlyLysArgGlyLysArgLysSerAspPheArgLysGlySerSerVal 96
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1065 CTGGGCAATACCGGAAAGGGGCAAAAGACTAGGCTCGAATTTCAAGTTTGAATCCAGGCTC 1124
Qy      97 GlnPheThrCysAengInaGlyTyrAspLeuGlnLysSerLysArgGlyIleThrCysMetLys 116
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1125 CAGTTTCACTGTCAACGAGGCTATGACTGCACAGGTTCCAAAGCGATCACTGTATGAA 1184
Qy      117 ValSerAspMetPheAlaAlaTyrSerAspHisArgProValCysArgAlaArgMetCys 136
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1185 GTGAGCACAATGTTGGGCTGGAGCGACACACAGCGCATCTCCGAGCCCGCATGTGT 1244
Qy      137 AspAlaHisLeuArgGlyProSerGlyIleIleThrSerProAsnPheProIleGlyTyr 156
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1245 GATGCCCACTTCAGAGCCCTCGGCGCATCTACCTCCCAATTTTCCCATTCAGATAT 1304
Qy      157 AspAsnAsnAlaHisCysValTyrIleIleThrAlaLeuAsnProSerLysValIleLys 176
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1305 GACAAACAATGCACCTGTGTGATCATCACAGCATCAACCCCTCCAAAGTATCAAG 1364
Qy      177 LeuAlaPheGluGluPheAspLeuGluArgGlyTyrAspThrLeuThrValGlyAspGly 196
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1365 CTCCCTTTGAGAGTTGATTTGAGAGGGGCTATGACACCTCGACGGTGGGATGGT 1424
Qy      197 GlnGlnAspGlyAspGlnLysThrValLeuTyrMetSerGlnAsnAlaCysSerAspSer 216
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1425 GGTACGAGTGGGAGCCGAAAGACAGTTCTTACATC----- 1460
Qy      217 ProIleThrProGlySerArgIleProGluSerMetSerGlyAspIleTyrArgGlnLys 236
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1461 -----CTGACAGGTACATGGTCCCGGATCTCATTTGCACACCAAT---CATCAATG 1511
Qy      237 TyrThrValLeuGluIleCysArgAspIleSerSerSerAspAlaArgSerGlySerVal 256
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1512 TGGCTCTCTCTCCAG-----ACTGATGGC---AGTGGAGCT--- 1544
Qy      257 ArgLysSerProLysThrSerAsnAlaValGluLeuValAlaProGlyThrGluIleGlu 276
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1545 -----TCCCTGGGATTCAGAGCTTCTTATGAAGAGATCGAG 1580
Qy      277 GlnGlySerCysGlyAspProGlyIleProAlaTyrGlyArgArgGluGlySerArgPhe 296
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1581 CAGGGCAGTTGCGGAGACCTTGGCATATCTGCATATGGCCGGAAGAGGCTCCCGGTTT 1640

```

```

Qy      297 HisIleGlyAspThrLeuLysPheGluCysGlnProAlaPheGluLeuValGlyGlnLys 316
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1641 CACCAAGGTACACACTCAAGTTTGAAGTCCAGAGCCCGCTTTAGCTGTGGGACAGAG 1700
Qy      317 AlaIleThrCysGlnLysAsnAengInaTyrSerAlaLysLysProGlyCysValPheSer 336
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1701 GCATCATCAGCCAAAAGAAATTAACAAATGGTCCGGCTTAAGAACGAGCTCGTGTCTCC 1760
Qy      337 CysPhePheAsnPheThrSerProSerGlyValValLeuSerProAsnTyrProGluAsp 356
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1761 TGGCTTCTCAACTTCACACAGCCCGTCTGGGGTTGTCTGTCTCCCACTCCACAGAGAC 1820
Qy      357 TyrGlyAsnHisLeuHisCysValTyrPheLeuIleLeuAlaArgProGluSerArgIleHis 376
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1821 TATGGCAACACCTTCACCTGTCGTGCTATCTGGCCAGGCTTGAGAGCCGATCCAC 1880
Qy      377 LeuAlaPheAsnAspIleAspValGluProGlnPheAspPheLeuValIleLysAspGly 396
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1881 CTGGCTTCAACGACATTAAGTGAAGCTCAAGTTCAGTTTATTTCCGTTCATCAAGATGGG 1940
Qy      397 AlaThrAlaGluAlaProValLeuGlyThrPheSerGlyAsnGlnLeuProSerSerIle 416
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1941 GCCACCCGCGAGGCGCCGCTCTGGGCACTTCACAGAAACGAGCTTCCCTCTCATC 2000
Qy      417 ThrSerSerGlyHisValAlaArgLeuGluPheGlnThrAspHisSerThrGlyLysArg 436
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2001 ACAAGCAGTGGCCACGAGGCGCGCTGAGTTCAGACTACCACTCCACAGGGAAGG 2060
Qy      437 GlyPheAsnIleThrPheThrThrPheArgHisAsnGluCysProAspProGlyValPro 456
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2061 GGCTTCAACATCACTTTTACCACTTCGACACAGAGAGCCCGGATCTCGGCGTTCCA 2120
Qy      457 ValAsnGlyLysArgPheGlyAspSerLeuGlnLeuGlySerSerIleSerPheLeuCys 476
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2121 GTAATATGGCAAAACGTTGGGAGACAGCTCAGCTGGGAGCTTCACATCTCTCTCTGT 2180
Qy      477 AspGlnGlyPheLeuGlyThrGlnGlySerGlnThrIleThrCysValIleLysGluGly 496
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2181 GATGAAGGCTTCTTGGGACTCAGGGCTCAGAGCCATCACTGCGTCTGAGAGAGAGGC 2240
Qy      497 SerValValTyrAsnSerAlaValLeuArgCysGlnAlaProCysGlyGlyHisLeuThr 516
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2241 AGCGTGTCTGAAACAGCGCTGTGCTGGGTGAACCTCCCTGTGGTCACTGACT 2300
Qy      517 SerProSerGlyThrIleLeuSerProGlyTyrProGlyPheTyrLysAspAlaLeuSer 536
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2301 TCGCCCAAGCGGACCATCTCTCTCGGCGCTGGCTGTCTTCAAGAGATGCTTGAAGC 2360
Qy      537 CysAlaThrValIleGluAlaGlnProGlyTyrProIleLysIleThrPheAspArgPhe 556
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2361 TGTGCTGGGTGATGAGGGCCGACGAGGCTACCCCAACCAACCTTCGACAGATTC 2420
Qy      557 LysThrGluValAsnTyrAspThrLeuGlnValArgAspGlyArgThrTyrSerAlaPro 576
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2421 AAAACCGAGGTCAACTATGACACCCCTGGAAGTACGCAATGGCGGCACTTATCGAGCC 2480
Qy      2481 TTGATCGGGGTTTACACAGGAGACCCAGGATTCACAGTCTCCATCACACAGCAATAC 2540
Qy      597 LeuTyrLeuLeuPheSerThrAspLysSerHisSerAspIleGlyPheGlnLeuArgTyr 616
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2541 CTCTAATCTCTCTTCTTACCGAACAAGTACCTCGGACATCGGCTTCAGCTCCGCTAT 2600
Qy      617 GluThrIleThrLeuGlnSerAspHisCysLeuAspProGlyIleProValAsnGlyGln 636
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2601 GAGACTATTAACCTGCAGCTGACAGCACTGTGATCCAGAAATCCAGTAAATGAGACAG 2660
Qy      637 ArgHisGlyAsnAspPheTyrValGlyAlaLeuValThrPheSerCysAspSerGlyTyr 656
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      2661 CGTCATGAGATGACTTCTACGTGGGCGCGCTGTGACCTTCAGCTGTGACTCGGCGTAC 2720

```

QY 657 ThrLeuSerAspGlyGluProLeuGluCysGluProAsnPheGlnTrpSerArgAlaLeu 676
 DB 2721 AATTATAGTAGAGGGAGAGCTTGGAGTGTAGGCCAACTTCCATGGAGCCGGGCTCTG 2780
 QY 677 ProSerCysGluAlaLeuCysGlyGlyPheIleGlnGlySerSerGlyThrIleLeuSer 696
 DB 2781 CCCATTGTGAAGCTCTGTGTGGCTTTCATTCAAGGCTCCAGTGGGACCATCTTGTGCG 2840
 QY 697 ProGlyPheProAspPheThrProAsnLeuAsnCysThrTrpIleIleGlnThrSer 716
 DB 2841 CCAGGATTCCTGACTTCTTACCCCAACAACCTTGAACCTGACCTGGATTATGAAACACTCT 2900
 QY 717 HisGlyGlyGlyValPhePheThrPheHisThrPheHisLeuGluSerGlyValAspTyr 736
 DB 2901 CATGGCAAGGGGTGTCTTCTTCCACACTTCCACTGGAAAGGGCCATGACTAC 2960
 QY 737 LeuLeuIleThrGluAsnGlySerPheThrGlnProLeuAsnArgGlnLeuThrGlySerArg 756
 DB 2961 CTCTCATCTACTGAAAGGCGAGCTTCCACCGAGCCCTGAGGCGACTAATCGATCTCGG 3020
 QY 757 LeuProAlaProIleSerAlaGlyLeuTyrGlyValAsnPheThrAlaGlnValArgPheIle 776
 DB 3021 CTGCCAGCTCCCATTCAGGCTGTGGCTCTATGGCAACTTCACTGCCAGGTCCTGCTTATC 3080
 QY 777 SerAspPheSerMetSerTyrGlyGlyPheAsnIleThrPheSerGlyTyrAspLeuGlu 796
 DB 3081 TCTGATTTCTCATGTCTATGAGGATTCAAACATCACTTCTCAGATGAGCATCTTGAG 3140
 QY 797 ProCysGlyGluGluProGlyGluValProAlaTyrSerIleArgGlyGlyLeuGlnPheGlyVal 816
 DB 3141 CCGGTGGAGAGCCCGAGGCTCCAGCCTCAAGCATCCGGAAGGGCTTGGAGTTGGGCTG 3200
 QY 817 GlyAspThrLeuThrPheSerCysPheProGlyTyrArgLeuGluGlyThrAlaArgIle 836
 DB 3201 GGGCAACCTTGAACCTTCTCTCTGCTCCCGGGATCCGTCTGAGAGGCAACCGCCGATC 3260
 QY 837 ThrCysLeuGlyGlyValArgArgArgLeuTrpSerSerProLeuProArgCysValAlaGlu 856
 DB 3261 ACGGCTCTGGGGGCGCAACGCGCTGTGGAGCTCGCTCGCCAAAGTGTGTGTGAG 3320
 QY 857 CysGlyAsnSerValThrGlyThrGlnGlyThrLeuLeuSerProAsnPheProValAsn 876
 DB 3321 TGTGGGAATTCAGTCAAGGCACTCAGGATCTTGTGTCTCCCAACTTCTCTGAGAC 3380
 QY 877 TyrAsnAsnAsnHisGlyCysIleTyrSerIleGlnThrGlnProGlyGlySerIleGln 896
 DB 3381 TACAATGAACATCATGATGATCTACTCATCCAGACCCAGCAAGGAAGGAATTCAG 3440
 QY 897 LeuValAlaArgAlaPheGlyLeuLeuSerGlyGlyAspValLeuValTyrAspGlyAsn 916
 DB 3441 CTGAAGCCCAAGGGATTCGAACCTCCGAAGAGAGATGCTCTCAAGGTTTATGATGGCAAC 3500
 QY 917 AsnAsnSerAlaArgLeuLeuGlyValPheSerHisSerGlnMetGlyValIleTrpLeu 936
 DB 3501 AACAACTCCGCCCTTGTCTGGAGTTTAAAGCATTTCTGAGATGATGGGGGAGACTTGG 3560
 QY 937 AsnSerThrSerSerSerLeuTrpLeuAspPheIleThrAspAlaGluAsnThrSerIle 956
 DB 3561 AACAGCAATCCAGAGCTGTGTGCTGATTTCATCATCGATCTGAAACACCAAGCAAG 3620
 QY 957 GlyPheGluLeuHisPheSerSerPheGlyLeuLeuIleCysGlyGlyAspProGlyThrPro 976
 DB 3621 GGGCTTGAACCTGACTTTCACACTTGAACCTCATGAAATGTGAGAGCCAGAGAACCCCC 3680
 QY 977 LysPheGlyTyrIleValHisAspGlyGlyValPheAlaGlySerSerValSerPheSer 996
 DB 3681 AAGTTGGCTACAAAGGTATCATGATGATGATGATTTTGCAGAGAGCTCCGTGCTTCCAC 3740
 QY 997 CysAspProGlyTyrSerLeuAsnArgGlySerGlyGlyLeuLeuCysLeuSerGlyGlyArg 1016
 DB 3741 TGTGACCTTGGAGTACAGCTCGCGGGTATGAGAGACTGCTGTGTGAGTGAAGAGCC 3800
 QY 1017 ArgThrTrpAspArgProLeuProThrCysValAlaGlyCysGlyGlyThrValArgGly 1036

DB 3801 CGAGCTGGGACCGGCTCTGCCCACTGTGTGCGCGAGTGTGAGAGGACAGTGAAGGA 3860
 QY 1037 GluValSerGlyGlnValLeuSerProGlyTyrProAlaProTyrGlyHisAsnLeuAsn 1056
 DB 3861 GAGGTGTCCGGGAGAGGTGTCTGTACCCGGGTATCCAGCTTCCATTAAACAATTCACAC 3920
 QY 1057 CysIleTrpThrIleGlnAlaGluAlaGlyCysThrIleGlyLeuHisPheLeuValPhe 1076
 DB 3921 TGCATCTGAGCACTCAAGACAGAGCGGCTGACCATTTGGGCTACATCTTCTGGTGT 3980
 QY 1077 AspThrGluGluValHisAspValLeuAsnGlyIleTrpAspGlyProValGluSerGlyVal 1096
 DB 3981 GACACAGAGAGGTATACACAGCTGTGCGCATCTGGAGGGCTGTGAGAGCGGGGCTT 4040
 QY 1097 LeuLeuLysGlyLeuSerGlyProAlaLeuProLysAspLeuHisSerThrPheAsnSer 1116
 DB 4041 CTGCTGAAGAGGTAGTGTGCGCGGCTGCCCAAGACCTGATAGCACTTCAACTCG 4100
 QY 1117 ValValLeuGlnPheSerThrAspPheThrSerLysGlnGlyPheAlaIleGlnPhe 1136
 DB 4101 GTCTCTGCACTTCAAGCACTGACTTCTTCAACCAAGACAGGCGCTTGCATTCAATT 4160
 QY 1137 SerValSerThrAlaThrSerCysAsnAspProGlyIleProGlyAsnGlySerArgSer 1156
 DB 4161 TCAGTGTCCACAGCAAGCTCTGTCATATGACCTGGGATTCGCAAGATGGAGTGGAGT 4220
 QY 1157 GlyAspSerTrpGluAlaGlyAspSerThrValPheGlnCysAspProGlyTyrAlaLeu 1176
 DB 4221 GGTGACAGTTGGGAAGCGGCGCATCCACAGTTCCTCAAGTGTGACCTTGCCATCCGCTG 4280
 QY 1177 GlnGlySerAlaGluIleSerCysValIleGlnAsnArgPhePheTrpGlnProSer 1196
 DB 4281 CAGGAAGTGCAGAGATCAGCTGTGTGAAGTCCGAAACAGGTTCCTTCTGGAGCCGACG 4340
 QY 1197 ProProThrCysIleAlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeu 1216
 DB 4341 CCGCAACATGATGTGCTCTCCGCGGGGAGACCTGACAGACATCTGAGATCAATCTTC 4400
 QY 1217 SerProAsnTyrProGluProTyrProProGlyGlyCysAspTrpLysValThrVal 1236
 DB 4401 TCACCAAAATTAACCGAACCCTTACCCGCGAGCAAGAGTGTGACTGAAAGTACCGTC 4460
 QY 1237 SerProAspTyrValIleAlaLeuValPheAsnIlePheAsnLeuGluProGlyTyrAsp 1256
 DB 4461 TCACCAAGACTAGTATGCTGCTGGTATTAACTTTAACTTTGAGACCTGGCTATATAC 4520
 QY 1257 PheLeuHisIleTyrAspGlyArgAspSerLeuSerProLeuIleGlySerPheTyrGly 1276
 DB 4521 TTCTCCATATCTAGAGAGCGGAGCTCTCAGCCCTCTCATAGGAAGCTTCTATATGC 4580
 QY 1277 SerGlnLeuProGlyArgIleGlySerSerSerAsnSerLeuPheLeuAlaPheArgSer 1296
 DB 4581 TCCAGCTCCCAAGCGGCAATTGAAGCAGCAGCAAGCCCTTCTCTCGCCTTCGCGAGC 4640
 QY 1297 AspAlaSerValSerAsnAlaGlyPheValIleAspTyrThrGluAsnProArgGluSer 1316
 DB 4641 GATGATGTGAGCAATGCTGCGCTTGTCTCATTTAATAGAAACCCCGGAGAGCA 4700
 QY 1317 CysPheAspProGlySerIleLysAsnGlyThrArgValGlySerAspLeuLysLeuGly 1336
 DB 4701 TGTTTGATCTGTGTTCCATCAAGACAGGCAAGGGGTGGGCTCCAGCTGAACTGAGGC 4760
 QY 1337 SerSerValThrTyrTyrCysHisGlyGlyTyrGluValGluGlyThrSerThrLeuSer 1356
 DB 4761 TCCTCCGACACTACTCTGCAAGGCGGCTTACCAAGTTAGGCACTTGACCTTGAGC 4820
 QY 1357 CysIleLeuGlyProAspGlyLysProValTrpAsnAsnProArgProValCysThrAla 1376
 DB 4821 TGCATCTGGGGCTTATGGAGCCCTGTGTGAACAATCCCGGCAAGTCTGACAGCC 4880
 QY 1377 ProCysGlyGlyGlnTyrValGlySerAspGlyValValLeuSerProAsnTyrProGln 1396

Db 4881 CCTGTGGGGACAGTATGTGGGTTCCGACGAGTGTCTTGTCCCACTACACCCAG 4940
 Qy 1397 AsnTyrThrSerGlyGlnIleCysLeuTyrPheValThrValProIleAspTyrValVal 1416
 Db 4941 AACACACCAAGTACAGATCTGCTGTATTTGTACTGTGCCCCAAGCAAGTATGTGTG 5000
 Qy 1417 PheGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspValValGluValHisAspGly 1436
 Db 5001 TTTGGCCAGTTCCGCTTCTTTCACACGCGCTTCAACACGCTGTGGAGGTTCAAGAGCC 5060
 Qy 1437 HisSerGlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThrGlyGluSerLeu 1456
 Db 5061 CACAGCAGACCTCGCGCTCTCTCAGCTCCCTCGGGCTCCCATACAGAGGAATCACTG 5120
 Qy 1457 ProLeuAlaThrSerArgGlnValLeuIleLysPheSerAlaLysGlyLeuAlaProAla 1476
 Db 5121 CCTTGGCCACCTCCCAATCAAGTTCATTAAATTCAAGCGCAAGGCTTCGACACGCC 5180
 Qy 1477 ArgGlyPheHisPheValTyrGlnAlaValProArgThrSerAlaThrGlnCysSerSer 1496
 Db 5181 AGAGGCTTCCACTTGTCTACCAAGCGGTTCTGAAACGACGCGCACGACGTGACGCTC 5240
 Qy 1497 ValProGluProArgTyrGlyLysArgLeuGlySerAspPheSerValGlyAlaIleVal 1516
 Db 5241 GTGCCGGAACCCCGCTATGGCAGAGGCTGGGCAGTGACTTCTCGTGGGGGCATCGTC 5300
 Qy 1517 ArgPheGlyCysAsnSerGlyTyrAlaLeuGlnGlySerProGluIleGluCysLeuPro 1536
 Db 5301 CGCTTCAATGCACATCCGGCTATGCCCTGCAGGGGTGCGCAGAGATCGAGTGCCTCCT 5360
 Qy 1537 ValProGlyAlaLeuAlaGlnThrAsnValSerAlaProThrCysValIleProCysGly 1556
 Db 5361 GTGCTGGGGCTTGGCCCAATGGAATGTTCTCAGCGCCACGCTGTGGTCCGTGGCA 5420
 Qy 1557 GlyAsnLeuThrGluArgArgGlyTyrHisLeuSerProGlyPheProGluProTyrLeu 1576
 Db 5421 GGCAACCTCACAGAGCCGAGGGGCAACCATCTGCTCCCTGGCTTCCCAAGACGCTACCTC 5480
 Qy 1577 AsnSerLeuAsnGlyValTyrLysIleValValProGluGlyValGlnIleGlnIleGln 1596
 Db 5481 AACGCTCAACCTGTGTGTGGAAGATCGTGTCCCCAGAGGCGCTGCATCCACATCAA 5540
 Qy 1597 ValValSerPheValThrGlnGlnAsnTyrAspSerLeuGluValPheAspGlyAlaAsp 1616
 Db 5541 GTTGTCACTTTTGTGACAGACAGAACTGGGACTCGCTGGAAGATTTGATGTGTCAAT 5600
 Qy 1617 AsnThrValThrMetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSer 1636
 Db 5601 AACACTGTAAACCATGCTGGGGAGTTTCTCAGGAACAACCGTGCCTGCTCTGAACAGC 5660
 Qy 1637 ThrSerArgGlnLeuTyrLeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPhe 1656
 Db 5661 ACCCTCAACCAAGCTCTACTCTTCACTTCTACAGATATCAGCTATGTGCGAGCTGCTC 5720
 Qy 1657 HisLeuGlnTyrLysThrValGlyLeuSerSerCysAspProIleProAlaValProSerAsn 1676
 Db 5721 CACTTGAAGTACAAACGAGTGGGCTGAGCAGTTGTCCGGAACCTGCTGTGCCAGATAC 5780
 Qy 1677 GlyValLysThrGlyGluArgTyrLeuValAsnAspValValSerPheGlnCysGluPro 1696
 Db 5781 GGGGTGAAGACTGGCGAGCGCTACTTGTGTGAATGATGTGTGTCTTTCAGATGTGAACCG 5840
 Qy 1697 GlyTyrAlaLeuGlnGlyHisAlaHisIleSerCysMetProGlyTyrThrValArgArgTyr 1716
 Db 5841 GGAATATCCCTCCACGGGCGACGCCACATCTCTGCAATGCCGGAACAGTGGCGCGATGG 5900
 Qy 1717 AsnTyrProProPheLeuCysIleAlaGlnCysGlyGlyThrValGluGlnMetGluGly 1736
 Db 5901 AACACCTCTCTCACTCTGTATGTGCACAGTGTGGGGAGACAGTGGAGAGATGAGAGGG 5960
 Qy 1737 ValIleLeuSerProGlyPheProGlyAsnTyrProSerAsnMetAspCysSerTyrLys 1756
 Db 5961 GTGATCTCTAGACCCCGGCTTCCAGGGCACTAACCCAGTAACTATGGAATGCTCTGGAAA 6020

Qy 1757 IleAlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGluPro 1776
 Db 6021 ATAGCACTGCCCTGGGCTTTGGAGCTCACATCAAGTCTCTGAATCTTCTCACCGAGCC 6080
 Qy 1777 AsnHisAspTyrIleGluIleArgAsnGlyProTyrGluThrSerArgMetGlyArg 1796
 Db 6081 AACCAAGACTACATPAGAAATCCGGAATGGCCCTATGAGACCAAGCCGATATGGGAAGA 6140
 Qy 1797 PheSerGlySerGluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTyr 1816
 Db 6141 TTCAGTGAAGCAGAGCTTCCAAAGCTCCCTCTCCACGTCGCCACGAGACACCGTGTAT 6200
 Qy 1817 PheHisSerAspHisSerGlnAsnArgProGlyPheLysLeuGlnTyrGln----- 1833
 Db 6201 TTCACAGCAGCACTTCCAGAAATCGGCAGAAATTCAGCTGGAGTATCAGATTGACT 6260
 Qy 1833 ----- 1833
 Db 6261 TACTCCACAGATTTCTCTCTCTGAGAGTTTGAATCTCTCGAGATTGAAGAAC 6320
 Qy 1834 -----Ala 1834
 Db 6321 AACCAACTCTCCCGCTCGCGCTTCTATGTCTGGAGTCTGATCTCTGCTTGTGAAGCC 6380
 Qy 1835 TyrGluLeuGlnGlyCysProAspProGluProPheAlaAsnGlyIleValArgGlyAla 1854
 Db 6381 TATGAACCTTAAGAGTGCACAGACCAAGACCCCTTGGCCATGTGAGGGAGGCT 6440
 Qy 1855 GlyTyrAsnValGlyGlnSerValThrPheGlyCysLeuProGlyTyrGlnLeuThrGly 1874
 Db 6441 GGCTACAACTGGGAGCAATCAGTACCTTGAAGTCTCCCGGGATATCAATTGACTGGC 6500
 Qy 1875 HisProValLeuThrCysGlnHisGlyTyrAsnArgAsnTyrAspHisProLeuProLys 1894
 Db 6501 CACCTGTCTCTCAAGTGTCAATGACCAACCGGAACCGGAGACCAACCCCTGCGCAAG 6560
 Qy 1895 CysGluValProCysGlyGlyAsnIleThrSerSerAsnGlnThrValTyrSerProGly 1914
 Db 6561 TGTGAAGTCCCTTGTGGGGGAACATCACTTTCACAGGCACTGTGTACTCCCGGGG 6620
 Qy 1915 PheProSerProTyrSerSerSerSerGlnAspCysValTyrPheIleThrValProIleGly 1934
 Db 6621 TTCCCTAGCCCGTATCTCACTCCAGAGACTGTGTCTGGGTATACCGTGCCATTGGC 6680
 Qy 1935 HisGlyValArgLeuAsnLeuSerSerLeuLeuGlnThrGluProSerGlyAspPheIleThr 1954
 Db 6681 CATGGCGTCCGCTCAACTCAGCTGTCTCAGACAGAGCCCTGTGAGATTTTCATCACC 6740
 Qy 1955 IleTyrAspGlyProGlnGlnThrAlaProArgLeuGlyValPheThrArgSerMetAla 1974
 Db 6741 ATCTGGATGGGCGCACAGCAACAGCACCGGCTCGGCTCTTTCACCCGAGCATGGC 6800
 Qy 1975 LysIleThrValGlnSerSerSerSerValLeuLeuLysPheHisArgAspAlaAla 1994
 Db 6801 AAGAAACAGTGCAGAGTTCATCCAACACAGTCTGTCAAGTTCACACCTGATGAGCC 6860
 Qy 1995 ThrGlyGlyIlePheAlaIleAlaPheSerAlaTyrProLeuThrLysCysProProPro 2014
 Db 6861 ACAGGGGGAGATCTTCCGCAATGCTTCTCCGTTATCACTACCAAAATGCCCTCTCC 6920
 Qy 2015 ThrIleLeuProAsnAlaGluValValThrGluAsnGluLysPheAsnIleGlyAspIle 2034
 Db 6921 ACCATCTCTCCCAACGCGGAAGTGTGCACAGAGAAATGAATTAATATGAGTGAATC 6980
 Qy 2035 ValArgTyrArgCysLeuProGlyPheThrLeuValGlyAsnGluIleLeuThrCysLys 2054
 Db 6981 GTAGCGTACAGATGCTCCCTGCTTACCTTGTGGGGAATGAATTTGACTGTGAAA 7040
 Qy 2055 LeuGlyThrTyrLeuGlnPheGluGlyProProProIleCysGluValHisCysProThr 2074
 Db 7041 CTGGAACTTACCTGAGTGTGAAGACCAACCCCGATATGTGAAGTGAAGTGTGCAACA 7100

Qy	2075	AsnGluLeuLeuThraSerThrGlyValIleLeuSerGlnSerTyrProGlySerTyr	2094
Db	7101	AATGACTTCTGACACACTCCACAGCGGATGCTTACGACCAAGCTACCTGGAAAGCTAT	7160
Qy	2095	ProGlnPheGlnThrCysSerTyrLeuValArgValGluProAspTyrAsnIleSerLeu	2114
Db	7161	CCCACTTCACAGACCTGCTCTGGCTGGAGAGTGAAGCCGCACTATACATCTCCCTC	7220
Qy	2115	ThrValGluTyrPheLeuSerGluuysGlnTyrAspGluPheGluIlePheAspGlyPro	2134
Db	7221	ACAGTGAATACTCTCTCCACGAGAAAGATATGATGAGTTGAAATTTTGAATGCTCA	7280
Qy	2135	SerGlyGlnSerProLeuLeuValAlaLeuSerGlyAsnTyrSerAlaProLeuIleVal	2154
Db	7281	TCAGGACAGAGTCTCTGCTGAAGCCCTCAGTGGAAATTACTACGCTCCCTGATTGTC	7340
Qy	2155	ThrSerSerAspAsnSerValTyrLeuArgTyrPheSerAspPheAlaTyrAsnArgLys	2174
Db	7341	ACCAAGCTCAAGCACTCTGTGTACCTGGTGGTATCTGATCAAGCTCAATCGGAAG	7400
Qy	2175	GlyPheLysIleArgTyrSerAlaProTyrCysSerLeuProArgAlaProLeuHisGly	2194
Db	7401	GGCTTCAGATCCGCTATTCAGCCCTTACTGACGCTCCCAAGGCTCCACTCCATGCG	7460
Qy	2195	PheIleLeuGlyGlnThrSerThrGlnProGlySerIleHisPheGlyCysAsnAla	2214
Db	7461	TTGATCTTAGGCGCACACGACCCAGCCGGGCTCATCCATTGGCTGCAAGGCC	7520
Qy	2215	GlyTyrArgLeuValGlyHisSerMetAlaIleCysThrArgHisProGlnGlyTyrHis	2234
Db	7521	GGCTACCCCTGGTGGGACACAGATGCGCATCTGTACCCGCGACCCCAAGGGCTACAC	7580
Qy	2235	LeuTyrSerGluAlaIleProLeuCysGlnAlaLeuSerCysGlyLeuProGluAlaPro	2254
Db	7581	CTGGAGACGGAAGCATCCCTCTCTGTCAGACTTTTCTGTGGCTTCTCTGAGCCGCC	7640
Qy	2255	LysAsnGlyMetValPheGlyLysGluTyrThrValGlyTyrLysAlaValTyrSerCys	2274
Db	7641	AAGAATGGAATGGTGTGGCAGAGATGACACATGGGAACCAAGGCGCATGTACAGTGC	7700
Qy	2275	SerGluGlyTyrHisLeuGlnAlaGlyValAlaThrAlaGluCysLeuAspThrGly	2294
Db	7701	AGTGAAGCTACCACTCCAGGCGAGCGCTGAGGCCCTGCAAGTGTCTTGGACACAGGC	7760
Qy	2295	LeuTyrSerAsnAlaGlnValProProGlnCysValProValThrCysProAspValSer	2314
Db	7761	CTATGGAGCAACCGCAATGTCGCCACAGTGTGCTGTGACTGTCTGTATGTCCAGT	7820
Qy	2315	SerIleSerValGluHisGlyArgTyrArgLeuIlePheGluThrGlnTyrGlnPheGln	2334
Db	7821	AGCATGAGCTGAGCATGGCCCATGGAGGCTTATCTTGAACACAGATACAGTGTCCAG	7880
Qy	2335	AlaGlnLeuMetLeuIleCysAspProGlyTyrTyrTyrThrGlyGlnArgValIleArg	2354
Db	7881	GCCCAAGCTGATGCTCATCTGTGACCTGGCTACTACTATACGGCCAAAGGCTATCCGC	7940
Qy	2355	CysGluAlaAsnGlyLysTyrPheSerLeuGlyAspSerThrProThrCysArgIleIleSer	2374
Db	7941	TGTCAAGCCCAATGCAATGAGACCTCGGGAGCTTACACGCCCACTGTCTGTACAGGG	8000
Qy	2375	CysGlyGluLeuProIleProProAsnGlyHisArgIleGlyThrLeuSerValTyrGly	2394
Db	8001	TGTGAGAGCTCCCGATTCGCCCATATGCGCACCGCATCGTAACACTGTCTGTACAGGG	8060
Qy	2395	AlaThrAlaIlePheSerCysAsnSerGlyTyrThrLeuValGlySerArgValArgGlu	2414
Db	8061	GCAACAGCACTCTTCTCTGCAATTCGGATACACACTGTGTGGGCTCCAGGGTGTGAG	8120
Qy	2415	CysMetAlaAsnGlyLeuTyrPheSerGlySerGlyValAlaArgCysLeuAlaGlyHisCysGly	2434
Db	8121	TGTATGGCCAAATGGCTCTGGAGTGGCTGTGAAGTCCGCTGTGGTGGACACTGTGGG	8180
Qy	2435	ThrProGluProIleValAlaAsnGlyHisIleAsnGlyGluAsnTyrSerTyrArgGlySer	2454
Db	8181	ACTCCGAGCCCATTTGCAACGAGACATCAATGGGAGAACTACAGCTACCGGGCAGT	8240
Qy	2455	ValValTyrGlnCysAsnAlaGlyPheArgLeuIleGlyMetSerValArgIleCysGln	2474
Db	8241	GTTGTGTACCAATGCAATGTGGCTTCCGCTGATCGGATGTCTGTGGCATCTCCAG	8300
Qy	2475	GlnAspHisHisTyrPheSerGlyLysThrProPheCysValProIleThrCysGlyHisPro	2494
Db	8301	CAGATCATCATCTGTGGGCAAGACCCCTTCTGTGTG-----	8339
Qy	2495	GlyAsnProValAlaAsnGlyLeuThrGlnGlyAsnGlnPheAsnLeuAsnAspValValLys	2514
Db	8339	-----	8339
Qy	2515	PheValCysAsnProGlyTyrMetAlaGluGlyAlaAlaArgSerGlnCysLeuAlaSer	2534
Db	8339	-----	8339
Qy	2535	GlyGlnTyrPheSerAspMetLeuProThrCysArgIleIleAsnCysThrAspProGlyHis	2554
Db	8339	-----	8339
Qy	2555	GlnGluAsnSerValArgGlnValHisAlaSerGlyProHisAspPheSerPheGlyThr	2574
Db	8339	-----	8339
Qy	2575	ThrValSerTyrArgCysAsnHisGlyPheTyrLeuLeuGlyThrProValLeuSerCys	2594
Db	8339	-----	8339
Qy	2595	GlnGlyAspGlyThrThrAspArgProArgProGlnCysLeuLeuValSerCysGlyHis	2614
Db	8340	-----CTGGTGTCTCTGGCCAT	8357
Qy	2615	ProGlySerProProHisSerGlnMetSerGlyAspSerTyrThrValGlyAlaValVal	2634
Db	8358	CCGGCTCCCGGCTCATCTCCCAATGTCTGGAACAGTTATCTGTGGAGCAGTGGTG	8417
Qy	2635	ArgTyrSerCysIleGlyLysArgThrLeuValGlyAsnSerThrArgMetCysGlyLeu	2654
Db	8418	CGGTACACCTCATCGGCACACGTACTGGTGGGAACAGCACCCGCACTGTGGGCTG	8477
Qy	2655	AspGlyHisTyrPheGlnTyrSerLeuProHisCysSerGlyThrSerValGlyValCysGly	2674
Db	8478	GATGACACTGAGCTGGCTCTCTCCACTGCTCACAGAAACCAAGCGGAGATTGGGCT	8537
Qy	2675	AspProGlyIleProAlaHisGlyIleArgLeuGlyAspSerPheAspProGlyThrVal	2694
Db	8538	GACCTGGAGTCCGGGCTCATGGCATCGTTTGGGGACAGCTTTGATCCAGGCACTGTG	8597
Qy	2695	MetArgPheSerCysGluAlaGlyHisValLeuArgGlySerSerGlyuArgThrCysGln	2714
Db	8598	ATGGCTTCAGCTGTGAAGTGGCCAGTGTCTCGGGGATCGTCAGAGCGCACTGTCAA	8657
Qy	2715	AlaAsnGlySerTyrPheSerGlySerGlnProGlyCysGlyValIleSerCysGlyAsnPro	2734
Db	8658	GCCAAATGGCTGTGGAGCGGCTCGACCTGATGTGAGATGATCTTGTGGAAACCTT	8717
Qy	2735	GlyThrProSerAsnAlaArgValAlaPheSerAspGlyLeuValPheSerSerSerIle	2754
Db	8718	GGGACTCCAAATATGCGCAGATTGTTCAGTATGAGCTGGCTTTTCTCCAGCTATAC	8777
Qy	2755	ValTyrGluCysArgGluGlyTyrTyrAlaThrGlyLeuLeuSerArgHisCysSerVal	2774
Db	8778	GTCATATAGTGGGGAAGATATCTACGCCACAGGCTCTCTCAGCGGTACGTCTGGTGC	8837
Qy	2775	AsnGlyThrTyrPheGlySerAspProGluCysLeuValIleAsnGlyAspProGly	2794
Db	8838	AATGTATCTGACAGAGCATGACCTGTAGTGTCTGTATAACTGTGTACCTTGGG	8897
Qy	2795	IleProAlaAsnGlyLeuArgLeuGlyAsnAspPheArgTyrAsnLysThrValThrTyr	2814

Db 8898 ATTCAGCCCAATGGCTTCGGCTGGGCAATGCTTCAAGTCAACAAACAGTGTGACTAT 8957
 Qy 2815 GlnCysVal1ProGly1YrMetMetGluSerHisArgValSerValLeuSerCysThrIys 2834
 Db 8958 CAGTGTCTCCCTGGCTTATGATGAGATCATGATGATATCTGTGGACTGCACACCAAG 9017
 Qy 2835 AspArgThrTrpAsnGlyThrIysProValCysValAlaLeuMetCysIysProProPro 2854
 Db 9018 GACCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9077
 Qy 2855 Leu1LeuProAsnGlyIysVal1ValGlySerAspPheMetTrpGlySerSerVal1ThrIyr 2874
 Db 9078 CTCATCCCAATGGGAAGTGTGGGTCTGACTTCAATGAGGGGTCAAGTGTGACTTAT 9137
 Qy 2875 AlaCysLeuGluGly1YrIrgInLeuSerLeuProAlaValPheThrCysGluGlyAsnGly 2894
 Db 9138 GCTGTGCTGAGGGGTACAGCTCTCCCTGCGGGGTGTTCACCTGTGAGGGGAAAGGG 9197
 Qy 2895 SerTrpThrGlyIrgInLeuProGlnCysPheProVal1PheCysGlyAspProGlyVal1Pro 2914
 Db 9198 TCCGTGACCGGAGAGCTGCTCACTGTTCCTGTGTGTGCGGGGATCCCTGGTGTCCG 9257
 Qy 2915 SerArgGlyIrgArgGluAspArgGlyPheSerIyrArgSerSerVal1SerPheSerCys 2934
 Db 9258 TCCGTGTGGAGGAG 9317
 Qy 2935 HisProProLeuVal1LeuVal1GlySerProArgArgPheCysGlnSerAspGly1YrThr 2954
 Db 9318 CATCCCTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 9377
 Qy 2955 SerGly1YrGlnProSerCys1LeuAspProThrLeuThrCysAlaAspProGly1Val1 2974
 Db 9378 AGTGGACCCAGCCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 9437
 Qy 2975 ProGlnPheGly1LeuAsnAsnSerGlnGly1YrGlnVal1GlySerThrVal1LeuPhe 2994
 Db 9438 CCACAGTTTGGGATACAGAACAAATCTCAGGGCTACAGAGTGGAGACACAGTCTCTTC 9497
 Qy 2995 ArgCysGlnIysGly1YrLeuLeuGlnGlySerThrIyrArgThrCysLeuProAsnLeu 3014
 Db 9498 CGTTGTCAAAAAGGCTACCTGCTTCAGGGCTCCACAGAGACTCTCCCAAACTGC 9557
 Qy 3015 ThrTrpSerGly1YrProProAspCysVal1ProHisCysArgGlnProGlnIyrPro 3034
 Db 9558 ACCGTGGGTGAAACCCAGCTGATGTGTCCCAACAGCTGACAGGACAGCAAGACCCCA 9617
 Qy 3035 ThrHisAlaAsnVal1GlyAlaLeuAspLeuProSerMetGly1YrThrLeuLeu-ThrPr 3054
 Db 9618 ACCGATGCCAAGCGGGGGGCTGTGATTTGGCTCCATGGGCTACAGCTCATCTACTCC 9677
 Qy 3054 AlaArgArgAlaSerProSerArgVal1AlaProSerThrAlaProAlaArgArgMetAl 3074
 Db 9678 TGGCAGAGAGGCTTCTCTCTCAAGGGTGTGCTCCAGACCGCACTCAAGAGCGAGATGC 9737
 Qy 3074 AlaGlyAlaSerArgProSerAlaTrpArgSerGly1ProVal1IysAspProSerTh 3094
 Db 9738 AGCTGGACAG 9797
 Qy 3094 rLeuProGly1SerHisArgSerProIysPro 3104
 Db 9798 ACTGCCCGGAG 9828
 RESULT 3
 AX685959 8010 bp DNA linear PAT 29-MAR-2003
 LOCUS Sequence 3 from Patent WO02064791.
 DEFINITION AX685959
 ACCESSION AX685959
 VERSION AX685959.1 GI:29371853
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Alsdbrook II, J.P., Anderson, D.W., Burgess, C.E., Boldog, F.L.,
 Casman, S.J., Colman, S.D., Edinger, S.R., Ellerman, K., Gerlach, V.,
 Gorman, L., Grosse, W.M., Guo, X., Herimann, J.L., Kexuda, R.,
 Lepley, D.M., Li, L., Macdougall, J.R., Millet, I., Pena, C.E.,
 Peyman, J.A., Rastelli, L., Rieger, D.K., Shinkets, R.A., Smithson, G.,
 Snyder, K.A., Stone, D.J., Tchernov, V.T., Vernet, C.A., Voss, E.Z.,
 Zerhuzen, B.D., Zhong, H., and Zhong, M.
 TITLE Proteins and nucleic acids encoding same
 JOURNAL Patent: WO 02064791-A 3 22-AUG-2002;
 Curagen Corporation (US)
 FEATURES
 source 1..8010
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 8010
 Score: 14142.50 Matches: 2605
 Percent Similarity: 97.71% Conservative: 1
 Best Local Similarity: 97.68% Mismatches: 2
 Query Match: 83.26% Indels: 59
 DB: 6 Gaps: 3
 US-10-016-248-2 (1-3104) x AX685959 (1-8010)
 Qy 1 MetAlaGlyVal1AspProProAlaLeuLeuLeuProCysSerIleSerAspCys 20
 Db 1 ATGGCGGGGGCCCCCTCCCGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 Qy 21 AlSerAsnGlnArgHisSerVal1GlyVal1GlyProSerGluLeuVal1IysGlnIle 40
 Db 61 GCTAGCATATGAGGACACTCCGTGGGGGTGAGGAGCCCTCCAGCTAGTCAAGAAACAATT 120
 Qy 41 GluLeuIysSerArgGlyVal1IysLeuMetProSerIyrAspAsnSerGlnIysThrSer 60
 Db 121 GAGTTGAAGTCTGAGGTGTGAAGCTGATGCCAGCAAGACAAACACACCAAGACCTCT 180
 Qy 61 ValLeuThrGlnVal1GlyVal1SerGlnIyrHisAsnMetCysProAspProGly1IlePro 80
 Db 181 GTGTTACTCAGAGTGGTGTGTGCCAAGACATATATATGTGTCAAGACCTGGCATGCC 240
 Qy 81 GluArgGly1YrAspLeuGlySerAspPheArgLeuGlySerSerVal1GlnPheThrCys 100
 Db 241 GAAAGGGGCAAAAGACTAGGCTCGGATTCAGGTAGGATCCAGCGCTCAGTTCACTGC 300
 Qy 101 AsnGluGly1YrAspLeuGlnIysSerIyrArgIleThrCysMetIysVal1SerAspMet 120
 Db 301 AACGAGGCTATATACCTGCAAGGAGTCCAAAGGATCACTGATGAAAGTGAAGGACATG 360
 Qy 121 PheAlaIatPheSerAspHisArgProVal1CysArgAlaArgMetCysAspAlaHisIleu 140
 Db 361 TTGGCGGCTGAGGACACACAGGCAAGCTGTGCGAGCCCGCATGTGTATGCCCACTT 420
 Qy 141 ArgGlyProSerGly1IleIleThrSerProAsnPheProIleGlnIyrAspAsnAla 160
 Db 421 CGAGGCCCCCTCGGAGCATCATCACTCCCAATTTCCCATTCAGTATGACAAACAAAGCA 480
 Qy 161 HisCysVal1TrpIleIleThrAlaLeuAsnProSerIyrVal1IleIysLeuAlaPheGlu 180
 Db 481 CACTGTGTGATCATCACAGCACTCAACCCCTCCAAAGTATCAAGCTGCTTGGAG 540
 Qy 181 GluPheAspLeuGluArgGly1YrAspThrLeuThrVal1GlyAspGlyIrgInAspGly 200
 Db 541 GAGTTGATTTGGAGAGGGGCTATGACACCTCAGCGGTGTGTGTGTGTGTGTGTGTGTGT 600
 Qy 201 AspGlnIyrThrVal1LeuIyrMetSerGlnAsnAlaCysSerAspSerProIleThrPro 220
 Db 601 GACCAGAAAGACAGTTCTCTCATATGTCTCAAAAGTCCGACAGTACAGCCCTTCACACCCA 660
 Qy 221 GlySerArgIleProGluSerMetSerGly1AspIleTrpArgGlnIysTrpThrVal1Leu 240

```
Db 661 GGCCTCCGATCCAGAGACATGCTGGGACATCTGAGGACAAATGACCTGACTT 720
Qy 241 GGUUUCyAArgApIleSerSerAspAlaArgSerGlySerValArgLysSerPro 260
Db 721 GAGACTGTGCGACATTAGCAGTTCAAGTGCAGGTTCAGTGAAGAGTCTCA 780
Qy 261 LysThrSerAspAlaValGluLeuValAlaProGlyThrGluIleGluGlnGlySerCys 280
Db 781 AAGACTTCTAATGCTGTGAACTTGTCTCTGGAGACAGATCGACAGGAGGAGTTGC 840
Qy 281 GlyAspProGlyIleProAlaTyrGlyArgArgGlnGlySerArgPheHisGlyAsp 300
Db 841 GGTGACCTGGCATCTGACATATGCGGAGGAGGAGGCTCCGGTTTCAACCAAGTGAC 900
Qy 301 ThrLeuLysPheGlnCysGlnProAlaPheGlnLeuValGlyGlnLysAlaIleThrCys 320
Db 901 AACTCAAGTTTGAAGTGCACCGCCCTTTGAGCTGTGGGAGCAAGAAAGGCATCATGC 960
Qy 321 GlnLysAsnAsnGlnTrpSerAlaLysLysProGlyCysValPheSerCysPhePheAsn 340
Db 961 CAAAGATTAACCATGCTCGGCTTAAGAGCCAGCTGCTGCTCTCTGCTTTTAC 1020
Qy 341 PheThrSerProSerGlyValValLeuSerProAsnTyrProGluAspTyrGlyAsnHis 360
Db 1021 TTCAACAGCCCGTGTGGGTTGTCTCTCTCCCAACTACCAAGAGCATATGGCAACAC 1080
Qy 361 LeuHisCysValTyrPheLysIleLeuAlaArgProGlnSerArgIleHisIleuAlaPheAsn 380
Db 1081 CTCACACTGTGCTGGCTCATCTTGGCCAGGCTTAAGAGCCGCACTCCAGCTGCTTCAAC 1140
Qy 381 AspIleAspValGluProGlnPheAspPheLeuValIleLysAspGlyAlaThrAlaGlu 400
Db 1141 GACATTACAGTGGAGCTCATGTTGATTCTCGGTATCAAGGATGGGCGCACCGCGAG 1200
Qy 401 AlaProValLeuGlyThrPheSerGlyAsnGlnLeuProSerSerIleThrSerSerGly 420
Db 1201 GCGCCCTCTGGGACCTTCTCGAGAAACAGATTCTCCCTCCATCACAGACAGTGGC 1260
Qy 421 HisValAlaArgLeuGlnPheGlnThrAspHisSerThrGlyLysArgGlyPheAsnIle 440
Db 1261 CACGTGACCGCTCGAGTTCCAGACTGACCACTCCAGAGGAGAGAGGGCTTCAACATC 1320
Qy 441 ThrPheThrThrPheAlaGlnIleAsnGlnCysProAspProGlyValProValAsnGlyLys 460
Db 1321 ACTTTTACCACTTCTGACCAACAGAGTGGCCGAGTCTGGCTTCACTTAATGCAAA 1380
Qy 461 ArgPheGlyAspSerLeuGlnLeuGlySerSerIleSerPheLeuCysAspGlnGlyPhe 480
Db 1381 CGGTTGGGAGACGCTCCAGCTGGGACGCTCCATCTCTCTGTGATGAAGGCTTC 1440
Qy 481 LeuGlyThrGlnGlySerGluThrIleThrCysValLeuLysGlnGlySerValValTrp 500
Db 1441 CTTGGAGCTCAGGGGCTCAGAGACATACCTGCTCTGAGAGGAGGAGCGGTGCTGG 1500
Qy 501 AsnSerAlaValLeuArgCysGlnAlaProCysGlyIleHisIleuThrSerProSerGly 520
Db 1501 AAGAGGCTGTGCTGGGTGTGAAGTCTCTGTGGTGTGCTCCTGAGCTTGGCCAGCGGC 1560
Qy 521 ThrIleLeuSerProGlyTyrTrpProGlyPheTyrLysAspAlaLeuSerCysAlaTrpVal 540
Db 1561 ACCATCTCTCTCCGGCTGGCTGCTTCTCAAGAGATGCTTGAAGCTGTGCTGGGTG 1620
Qy 541 IleGluAlaGlnProGlyTyrProIleLysIleThrPheAspArgPheLysThrGlnVal 560
Db 1621 ATTGAAGCCCAAGCAGGCTAACCCCAATCAAAATCACCTTGCAGACGATTCAAAACCAAGGTC 1680
Qy 561 AsnTyrAspThrLeuGlnValAlaArgAspGlyArgThrTyrSerAlaProLeuIleGlyVal 580
Db 1681 AACTATGACACCTGGAATACGCAATGGCGGAGCTTACTACGCCCCCTTGTATGGGGTT 1740
Qy 581 TyrHisGlyThrGlnValProGlnPheLeuIleSerThrSerAsnTyrLeuTyrLeuLeu 600
1741 TACCAAGGAGCCAGGTTCCCGAGTTCTCATCAGACCAAGCAATCACTTACCTCTC 1800
Qy 601 PheSerThrAspLysSerHisSerAspIleGlyPheGlnLeuArgTyrGluThrIleThr 620
Db 1801 TTCTTCAAGCAAGATGATCTGGACATCGGCTTCCACTCCGCTATATAGACTAATACA 1860
Qy 621 LeuGlnSerAspHisCysLeuAspProGlyIleProValAsnGlyGlnArgHisGlyAsn 640
Db 1861 CTGCACTCAGACCACTGTCTGGATCCAGGAATCCAGTAATAGCAAGGCTCATGGAAAT 1920
Qy 641 AspPheTyrValGlyAlaLeuValThrPheSerCysAspSerGlyTyrThrLeuSerAsp 660
Db 1921 GACTTCTAGTGGGCGGCTGTGTGACCTTCACTGTGACTCGGCTTACATTAAGTGAC 1980
Qy 661 GlyGluProLeuGlnCysGluProAsnPheGlnTrpSerArgAlaLeuProSerCysGlu 680
Db 1981 GGGAGGCTCTGAGAGTGAAGCCCACTTCCAGTGAAGCCGAGGCTTGGCCAGTTGTGA 2040
Qy 681 AlaLeuCysGlyGlyPheIleGlnGlySerSerGlyThrIleLeuSerProGlyPhePro 700
Db 2041 GCTCTGTGTGGCTTCAATCAAGGCTCAAGTGGAGCCATCTTGTCCGAGGGTTCCCT 2100
Qy 701 AspPheTyrProAsnAsnLeuAsnCysThrTrpIleIleGluThrSerHisGlyLysGly 720
Db 2101 GACTTCTACCCCAACACTTGAATGACACTGAGATTATCGAAACATCTCATGGCAAGGT 2160
Qy 721 ValPhePheThrPheHisThrPheHisLeuGlnSerGlyIleHisAspTyrLeuLeuIleThr 740
Db 2161 GTGTCTTCACTTCCACACCTTCCAGCTGGAAAGTGGCCATGACTTACTCTCATCACT 2220
Qy 741 GlnAsnGlySerPheThrGlnProLeuArgIleuThrGlySerArgLeuProAlaPro 760
Db 2221 GAGAAAGGAGATTACCAACCCCTGAGGAGCTTAATGTGATCTCGGCTCCAGCTCCC 2280
Qy 761 IleSerAlaGlyLeuTyrGlyAsnPheThrAlaGlnValArgPheIleSerAspPheSer 780
Db 2281 ATCAGCGCTGGGCTCTAAGGCACTTCACTGCCAGGTCCGCTTCACTCTGATTTCTCC 2340
Qy 781 MetSerTyrGlnGlyPheAsnIleThrPheSerGluTyrAspLeuGluProCysGlnGlu 800
Db 2341 ATGTCAATGAAGATTCATCATCACTTCCAGAGTACGACTTGGAGCCCTGAGAGAG 2400
Qy 801 ProGluValProAlaTyrSerIleArgLysGlyLeuGlnPheGlyValGlyAspThrLeu 820
Db 2401 CCCAGAGTCCAGCTCAGCATCCGAAAGGCTTGCAGTTGGCGTGGGAGAACCTTGG 2460
Qy 821 ThrPheSerCysPheProGlyTyrArgLeuGlnGlyThrAlaArgIleThrCysLeuGly 840
Db 2461 ACCTTCTCTGCTTCCCGGGGTACCGTGTGAGGGACCGCCGCAATCAGTGCCTGGGG 2520
Qy 841 GlyArgArgArgLeuTrpSerSerProLeuProArgCysValAlaGluCysGlyAsnSer 860
Db 2521 GGCAGAGCGGCTGTGAGCTGAGCTCTCTCCAAAGGTGTGTCTGTGATGGGAATTCA 2580
Qy 861 ValThrGlyThrGlnGlyThrLeuLeuSerProAsnPheProValAsnTyrAsnAsnAsn 880
Db 2581 GTCAAGGACTCAGGGTACTTTGCTGTCTGCCCACTTCTGTGAATCAATTAACAT 2640
Qy 881 HisGlnCysIleTyrSerIleGlnThrGlnProGlyLysGlyIleGlnLeuLysAlaArg 900
Db 2641 CATAAATGCAATCACTCATCCAGACCCAGCCAGGGAAGGGAATTCAGTGAAGACAGG 2700
Qy 901 AlaPheGlnLeuSerGlnGlyAspValLeuLysValTyrAspGlyLysAsnAsnAsnAla 920
Db 2701 GCATTGGAATCTCCAGAGAGATGTCTCAAGTTTAATGATGGCAACAACTCCGCC 2760
Qy 921 ArgLeuLeuGlyValPheSerHisSerGlnLysMetGlyValThrLeuAsnSerThrSer 940
Db 2761 CGTTTGTGGAGATTTTTATGCAATCTGAGATGATGGGGGTGACTTGAACAGCAATCC 2820
Qy 941 SerSerLeuTrpLeuAspPheIleThrAspAlaGluAsnThrSerTyrGlyPheGluLeu 960
Db 2821 AGCAGTCTGTGGCTTATTTCACTGATGTGAACCAACAGCAAGGAGCTTTGAACTTG 2880
```

Qy	961	HisPheSerSerPheGluLeuIleValCysGluAspProGlyThrProIlePheGluThr	980	Qy	1321	GlySerIleValAsnGlyThrArgValGlySerAspLeuIleGlySerSerValThr	1340
Db	2881	CACTTTCACACTTGAATCATCAAAATGTAAGAACCCAGAACCTTGGCTAC	2940	Db	3961	GGTTCCATCAAGAACCGGACACCGGCTGGGTTCCGACTGTAAGCTGGCTCTCCGTAC	4020
Qy	981	LysValHisAspGluGlyHisPheAlaGlySerSerValSerPheSerCysAspProGly	1000	Qy	1341	TyrTyrCysHisGlyGlyTyrGluValGlyGlyThrSerThrLeuSerCysIleGly	1360
Db	2941	AAGTTTCATATGAAGGTCATTTTGCAGGGAGCTCCGTCCTTCACTGTGACCTCGGA	3000	Db	4021	TACTACTGCCACGGGGGCTACGAAGTTGAGGGACCTCGACCTGAGCTCATCTGGGG	4080
Qy	1001	TyrSerLeuArgGlySerGlyLeuLeuLeuCysLeuSerGlyGluArgGlyThrTrpAsp	1020	Qy	1361	ProAspGlyLysProValTrpAsnAsnProArgProValCysThrLysAspCysGly	1380
Db	3001	TACAGCTCGCGGGTAGAGAGAGCTGCTGTGTGAGTGAGAGCCCGACCTGGAGAC	3060	Db	4081	CCTGATGGGAAGCCCGTGTGAACAATCCCGGACAGTGCACAGCCCTGTGGGGA	4140
Qy	1021	ArgProLeuProThrCysValAlaGlyCysGlyValThrValArgGlyGluValSerGly	1040	Qy	1381	GlnTyrValGlySerAspGlyValValLeuSerProAsnTyrProGlnAsnTyrThrSer	1400
Db	3061	CGGCTCTGCCCACCTGTGTGCGAGGTGAGAGAGACAGTGAAGAGAGAGTTCGAGG	3120	Db	4141	CAGTATGTGGGTTTCGAGAGAGAGTGTCTTCCCGCACTACCCCGAATACACCACT	4200
Qy	1041	GlnValLeuSerProGlyTyrProAlaProTyrGlyLysHisAsnLeuAsnCysIleTrpThr	1060	Qy	1401	GlyGlnIleCysLeuTyrPheValThrValProLysAspTyrValValPheGlyGlnPhe	1420
Db	3121	CAGGTGCTGTACCCGGGTATCCAGCTCCCTATGAACAATCTCACTGTGATCGAAC	3180	Db	4201	GGAAGATCTGCTTGTATTTTGTACTGTGCGCAAGACTATGTGTGTGGCCAGTTC	4260
Qy	1061	IleGluAlaGluAlaGlyCysThrIleGlyLeuHisPheLeuValPheAspThrGlyGlu	1080	Qy	1421	AlaPhePheHisThrAlaLeuAsnAspValValGluValHisAspGlyHisSerGlnHis	1440
Db	3181	ATCGAACAGAGCGCGCTGCACCATTTGGGCTACATTCCTGCTGTGACACAGAGAG	3240	Db	4261	GCCTTCTTTACACGGGCTCAACGAGTGTGAGGTTTCAACAGCGCCACAGCAGAC	4320
Qy	1081	ValHisAspValLeuArgIleTrpAspGlyProValGlySerGlyValLeuLeuLysGlu	1100	Qy	1441	SerArgLeuLeuSerSerLeuSerGlySerHisThrGlyGlySerLeuProLeuAlaThr	1460
Db	3241	GTTCACACACGTGCTGGCCATCTGGAGATGGGCTGTGAGAGCGGGGTTCGTGAAGAG	3300	Db	4321	TCCGGCTCTTCAGCTCCCTCTCGGGCTCCCATACAGAGAAATCATGCGCTTGGCCACC	4380
Qy	1101	LeuSerGlyProAlaLeuProLysAspLeuHisSerThrPheAsnSerValValLeuGln	1120	Qy	1461	SerAsnGlnValLeuIleLysPheSerAlaLysGlyLeuAlaProAlaArgGlyPheHis	1480
Db	3301	CTGATGGCCCGGCGCTGCCAAGACCTGCATAGCACTTCACTCGTGTCTCTGACG	3360	Db	4381	TCCATCAAGTTCTCATTAATTCAGGCGCAAGAGGCTGCGACAGCAGCAGGCTTCAC	4440
Qy	1121	PheSerThrAspPhePheThrSerLysGlnGlyPheAlaIleGlnPheSerValSerThr	1140	Qy	1481	PheValTyrGlnAlaValProArgThrSerAlaThrGlnCysSerSerValProGluPro	1500
Db	3361	TTTCAGCATGACTCTTCACACAGACAGAGGCTTTCGCAATTCATTTTCAGTTCACA	3420	Db	4441	TTTGTCTACCAACGGTCTCTTCCACACGAGCCACGAGGACGCTGTGCGGGAACC	4500
Qy	1141	AlaPheSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerGlyAspSerTyr	1160	Qy	1501	ArgTyrGlyLysArgLeuGlySerAspPheSerValGlyValAlaIleValArgPheGluCys	1520
Db	3421	GCAACGTCTGTCAATGACCTCGGATCCCGAGATGGAGTCCGAGGTGGAGAGTTGG	3480	Db	4501	CGCTATGGCAAGAGCTGGGCACTTTCGCTGGGGGCAATCGTCCCTTCGAAATGC	4560
Qy	1161	GluAlaGlyAspSerThrValPheGlnCysAspProGlyTyrAlaLeuGlnLysAla	1180	Qy	1521	AsnSerGlyTyrAlaLeuGlnGlySerProGluIleGluCysLeuProValProGlyAla	1540
Db	3481	GAAACCGGCACTCCACAGTCTTCAGTGTACCTTGGCTACGGCTGCGAGGAAGTCA	3540	Db	4561	AACCTCGGCTATCCCTCGACAGGGGTGCGCAGAAATGAGTGCCTCCCTGTGGGGGCC	4620
Qy	1181	GluIleSerCysValLysIleGluAsnArgPhePheTrpGlnProSerProProThrCys	1200	Qy	1541	LeuAlaGlnTrpAsnValSerAlaProThrCysValAlaProCysGlyLysAsnLeuThr	1560
Db	3541	GAGATCAGCTGTGGAAGATCAGAAACAGGTTCTTTCGAGAGCCAGCCGCCAACATGC	3600	Db	4621	TTGGCCCAATGAATGCTCAGCGCCACGCTGTGTGCTCGGTGTGAGGCAACTCACA	4680
Qy	1201	IleAlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeuSerProAsnTyr	1220	Qy	1561	GluArgArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeuAsn	1580
Db	3601	ATCCTCTCTCGGGGGAGACCTTCAAGACCATCTGAGTCACTCTCCCAAAATTAC	3660	Db	4681	GAGCGAGGGGCAACATCTGTCTCCCTGGCTTCCAGAGCCGTACCTCAACAGCTTCAC	4740
Qy	1221	ProGluProTyrProProGlyLysGlyCysAspTrpLysValThrValSerProAspTyr	1240	Qy	1581	CysValTrpLysIleValValProGluGlyAlaGlyIleGlnIleGlnValValSerPhe	1600
Db	3661	CCAGAACCTTACCCGCGAGAGAGAGTGTACTGGAAAGTACCGTCTCACCGAGTAC	3720	Db	4741	TGTGTGTGAAGATCGTGTCCCGAAGAGGCTGGATCCAGATCCAAAGTTTGCAGTTT	4800
Qy	1241	ValIleAlaLeuValPheAsnIlePheAsnLeuGluProGlyTyrAspPheLeuHisIle	1260	Qy	1601	ValThrGluGlnAsnTrpAspSerLeuGluValPheAspGlyValAspAsnThrValThr	1620
Db	3721	GTCATCGCCCTGATTTTAACTTTAACTTGAAGCTGTGAGCTGTGACTTCTCATATC	3780	Db	4801	GTCACAGAGCAGAACTGGGACTCGCTGGAAGTATTTGATGTGCAGATTAACTGTAAAC	4860
Qy	1261	TyrArgGlyArgAspSerLeuSerProLeuIleGlySerPheTyrGlySerGlnLeuPro	1280	Qy	1621	MetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerThrSerAsnGln	1640
Db	3781	TACACGAGACGGAGCTCTCTACGCCCTCTCATAGAAAGCTTATGTGCTCCACACTCCA	3840	Db	4861	ATGCTGGAGGATTTCTCAGGAACAACCGTCCGCTTGTGAACAGACCTCCACAG	4920
Qy	1281	GlyArgIleGlySerSerSerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerVal	1300	Qy	1641	LeuTyrLeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHisLeuGlyTyr	1660
Db	3841	GGCGGCATTTGAACAGACAGACAGCTTCTTCTCGCTTCCGACGAGTGCATCTGTG	3900	Db	4921	CTTCACTTATTTCTAATCAGATATCAGATATCAGACTGCTTCCACTTGGAGTAC	4980
Qy	1301	SerAsnAlaGlyPheValIleAspTyrThrGluAsnProArgGlySerCysPheAspPro	1320	Qy	1661	LysThrValGlyLeuSerSerCysProGluProAlaValProSerAsnGlyValLysThr	1680
Db	3901	AGCAATGCTGGCTTCGTCACTTGAATATACAGAAACCGCGGAGTATGTTTTCATCT	3960	Db	4981	AAAACGGTGGCTGAGCAATTTGTCGGAACCTGCTGTGCCAGTAAACGGGGTGAAGACT	5040
Qy				Qy	1681	GlyGluArgTyrThrLeuValAsnAspValValSerPheGlnCysGluProGlyTyrAlaLeu	1700

```

Db      5041 GGGAGCGCTACTGTGATGATGTGTCTTCCAGTGTGAGCCGGGAATTCCTC 5100
Qy      1701 GINGLHIALHIALIleSerCysMetProGlyThrValArgArgTyrAsnTyrProPro 1720
Db      5101 CAGGGCCAGCCCACTCTCTGATGATCCCGGAAGAGTGGCGGATGAGACTTACCTCTCT 5160
Qy      1721 ProLeuCyS1leAlaGlnCysGlyThrValGluGluMetGluGlyValIleLeuSer 1740
Db      5161 CCACCTCTGTATTGACACAGTGTGGGGGAACAGTGAAGAGATGAGAGGGGTATCTTAGC 5220
Qy      1741 ProGlyPheProGlyAsnTyrProSerAsnMetAspCysSerTrpValIleAlaLeuPro 1760
Db      5221 CCCGGCTTCCAGGCAACTACCCCACTAATGAGCTCTCTCTGGAAAAATAGACTGCC 5280
Qy      1761 ValGlyPheGlyValAlaIleGlnPheLeuAsnPheSerThrGluProAsnHisAspTyr 1780
Db      5281 GTGGGCTTTGAGCTCACTACATCAGTCTCTGAACCTTCTCACCGAGCCCAACCAAGACTAC 5340
Qy      1781 IleGluIleArgAsnGlyProTyrGluThrSerArgMetGlyArgPheSerGlySer 1800
Db      5341 ATAGAAATCCGGAATGGCCCTATGAGACAGCCGATGATGGGAATTCAGTGAAGC 5400
Qy      1801 GluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTyrPheHisSerAsp 1820
Db      5401 GAGCTTCCAGAGCTCCCTCTCTCTCCACGTCACAGACACACCGTGTATTTCCACAGGAC 5460
Qy      1821 HisSerGlnAsnArgProGlyPheLeuGluGluTyrGlnAlaTyrGluLeuGlnGluCys 1840
Db      5461 CACTCCCAAGATCCGCGCAAGATTCAGCTGAGATTCAGGCTTGAACCTTCAAGAGTGC 5520
Qy      1841 ProAspProGluProPheAlaAsnGlyIleValArgGlyAlaGlyTyrAsnValGlyGln 1860
Db      5521 CCAGACCCAGAGCCCTTTGCAATGGCATGTGTGAGGGAGAGCTGCTCAACGTGGGCA 5580
Qy      1861 SerValThrPheGluCysLeuProGlyTyrGlnLeuThrGlyHisProValLeuThrCys 1880
Db      5581 TCAGTGAACCTTCAGAGTGCCTCCCGGGTATTAATTGACTGGCCACCTGTCTCAGCTGT 5640
Qy      1881 GlnHisGlyThrAsnArgAsnTyrAspHisProLeuProLysCysGluValProCysGly 1900
Db      5641 CAAATGGCACCAACCGGAATCGGACCAACCCCTGCCCCAAGTGAAGTCCCTTGTGGC 5700
Qy      1901 GlnAsnIleThrSerSerAsnGlyThrValTyrSerProGlyPheProSerProTyrSer 1920
Db      5701 GGGAACTCACTTCTTCCAAAGGACGTGTACTCCCGGGGTCTCCCTACCCCGTACTCC 5760
Qy      1921 SerSerGlnAspCysValTyrPheIleThrValProIleGlyHisGlyValArgLeuAsn 1940
Db      5761 AGCTCCAGAGACTGTGTCTGGCTGATCAACCTGCCATTTGGCCATGGCGTCCGCTCAC 5820
Qy      1941 LeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTyrAspGlyProGln 1960
Db      5821 CTGAGCTGTGTGAGACAGAGCCCTCTGAGATTTCACATCTGGATGGGACAG 5880
Qy      1961 GlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaIleArgThrValGlnSer 1980
Db      5881 CAHAAGCACACAGGCTCGGCGCTTTCACCCCGAGCATGGCCAAAGAAACAGTGAAGT 5940
Qy      1981 SerSerAsnGlnValLeuLeuLysPheHisArgAspAlaAlaThrGlyIlePheAla 2000
Db      5941 TCATCCAAACAGGCTCGCTCAAGTTCACCGTGAAGCAGCCACAGGGGGATCTTCCGC 6000
Qy      2001 IleAlaPheSerAlaTyrProLeuThrLysCysProProProThrIleLeuProAsnAla 2020
Db      6001 ATAGCTTCTTCCTGATTCACATCCACCAATGCTCTCCCAACCTCTCCCAACGCGC 6060
Qy      2021 GluValValThrGluAsnGluGluPheAsnIleGlyAspIleValArgTyrArgCysLeu 2040
Db      6061 GAAGTGTGTCAAGATGAATGAATTCATATAGGACATCGTAGAGCTACAGAGCTTC 6120
Qy      2041 ProGlyPheThrLeuValGlyAsnGluIleLeuThrCysLysLeuGlyThrTyrLeuGln 2060

Db      6121 CTGGCTTACTCTTAGTGGGAAATGAATTCAGACTGCAAACTTGGAACTTACCTGAG 6180
Qy      2061 PheGluGlyProProProIleCysGluValHisCysProThrAsnGluLeuLeuThrAsp 2080
Db      6181 TTGGAAGACACACCCCGATATGTGAATGCACTGTCCAAATAATAGCTTCTGACAGAC 6240
Qy      2081 SerThrGlyValIleLeuSerGlnSerTyrProGlySerTyrProGlnPheGlnThrCys 2100
Db      6241 TCCACAGGGGTATCTGTAGCCAGAGCTACCTGTGAAGCTATCCAGTTCAGAGCTTC 6300
Qy      2101 SerTrpLeuValArgValGluProAspTyrAsnIleSerLeuThrValGluTyrPheLeu 2120
Db      6301 TCTTGCTGTGTAGAGTGAAGCCCACTATACATCTCCCTACAGTGAAGTACTTCTTC 6360
Qy      2121 SerGluLysGlnTyrAspGluPheGluIlePheAspGlyProSerGlyGlnSerProLeu 2140
Db      6361 AGGAGAACCAATATGATGATTTGAGATTTTGTATGTGTCATCAGACAGAGTCTCTCG 6420
Qy      2141 LeuLysAlaLeuSerGlyAsnTyrSerAlaProLeuIleValThrSerSerAsnSer 2160
Db      6421 CTGAAGCCCTCAGTGGGAATTACTAGTCCCTGATGTGATCCAGCTCAAGCAACTCT 6480
Qy      2161 ValTyrLeuArgTyrSerSerAspHisAlaTyrAsnArgLysGlyPheLysIleArgTyr 2180
Db      6481 GTGTACTGCTGTGTGATCTGATCAGCCCTTACATCGGAAGGGCTTCAAGATCCGCTAT 6540
Qy      2181 SerAlaProTyrCysSerLeuProArgAlaProLeuHisGlyPheIleLeuGlyGlnThr 2200
Db      6541 TGAAGCCCTTACTGTGAGCTGTGCGCAGGGCTCACCTCCATGCTTCAATGAGCCAGAC 6600
Qy      2201 SerThrGlnProGlyGlySerIleHisPheGlyCysAsnAlaGlyTyrArgLeuValGly 2220
Db      6601 AGCAACCAAGCCCGGGGGCTTCATCACTTGTGGTCAACAGCCGGCTTACCGCTGGGGA 6660
Qy      2221 HisSerMetAlaIleCysThrArgHisProGlnGlyTyrHisLeuTyrSerGluAlaIle 2240
Db      6661 CACACATGGCGCATCTGTATCCCGGACCCCGAGGGCTTACCACTGTGAGGAGAACCATC 6720
Qy      2241 ProLeuCysGlnAlaLeuSerCysGlyLeuProGluAlaProLysAsnGlyMetValPhe 2260
Db      6721 CCTCTCTGAAGCTCTTCTGTGGGCTTCTTAAGGCCCCCAAGAAATGATGTGTT 6780
Qy      2261 GlyLysGluTyrThrValGlyThrLysAlaValTyrSerCysSerGluGlyTyrHisLeu 2280
Db      6781 GCGAAGAGTACACAGTGGGAACCAAGGCCATGACAGCTGCAAGGCTTACCACTTC 6840
Qy      2281 GlnAlaGlyValGluAlaThrAlaGluCysLeuAspThrGlyLeuTyrSerAsnArgAsn 2300
Db      6841 CAGGAGGCGCTGAGGCCACTGACAGAGTGTGTGACACAGGCTTATGAGCAACCGCAAT 6900
Qy      2301 ValProProGlnCysVal-----ProValThrCysPro 2311
Db      6901 GTCCACCAACAGTGTGTCCGAGATCTCGGGCAATGAGCGGGGTGTGACTGTGCT 6960
Qy      2312 AspValSerSerIleSerValGluHisGlyArgTyrArgLeuIlePheGluThrGlnTyr 2331
Db      6961 GATGTCAGTAGCATCAGCTGAGACATGAGCGCATGAGGCTTATCTTTGAGACACAGTAT 7020
Qy      2332 GlnPheGlnAlaGlnLeuMetLeuIleCysAspProGlyTyrTyrTyrThrGlyGlnArg 2351
Db      7021 CAGTTCAGAGGCCAGCTGATGCTCATCTGTGACCTGTGCTACATATATGCGCAAGG 7080
Qy      2352 ValIleArgCysGlnAlaAsnGlyLysTrpSerLeuGlyAspSerThrProThrCysArg 2371
Db      7081 GTCATCCGCTGTAGAGCCCAATGGAATGAGAGCTCGGGAGCTTACGCCCACTGCGCA 7140
Qy      2372 IleIleSerCysGlyGluLeuProIleProProAsnGlyHisArgIleGlyThrLeuSer 2391
Db      7141 ATCATCTCCGTGTGAGAGCTCCGATTTCCCAATGGCCACCGCATCGGAACATGTGT 7200
Qy      2392 ValTyrGlyAlaThrAlaIlePheSerCysAsnSerGlyTyrThrLeuValGlySerArg 2411
Db      7201 GTCTAGGGGCAACAGCATCTTCTCTGCAATTCGGAATACACACTGTGGGCTCCAGG 7260

```


misc_feature	17458375631ANG1ANG658FRC1ACG5CPD1YVNH1.5GDGFSYRDTVVYQCNPGF BLVGSVAC1CQDHRKMGQTFVCP1TICGPNPAPGFTNGSEFANLVNAPNTCNTRY LIQGSVRQCKNSGQWSSPLP7CQRVNC5DPGFVENA1RHQONFPESEFTEYM1LYX CKKGHLLGNSALTCVANGDLMBSLPCFKAL1SCGHFVQANAVL1TGE1FTGAVVHS CRGSJLL1GNDTRVCOBDSHMSGALPHICGNNGFCGDGTPAHGSR1GDDFPKTKLL RFSCEMGHQLRCSPERTCLANGSMG1OPCEAVSGCNGNPPTNCTM1YRSNG1LFPSS V1YACMBEYKTSGLMTRHCTANG1MTG1TAPDC1T1SCDPR1LANG1OPG1D1PFTNK V5OONPCBYNVA1TSA1TR1CTYDGRKMNRSKVCYCAV1LCPDPVON1TNG5SDFRNG S1S1SCMDG1QULSHS1LS1CBGRVYKWS1PQ1C1D1PACBP1LPAEGR1SGKSP1YK SEVPFOCKSP1T1GVSSRRVCOAD1GTW51OP1C1D1PAPN1CPCPGTPHFG1QWS5RG LEV5GTV1FVCRKGYH1QGS1TTR1CLANT1W5G1OTEC1P1ACR1OPETPAHAD1VAP1D LPTG1YTL1VYRCHGPF1LAG58HRTCKAD1MKTGSKSPVCSKGV1RENE1VTKTP1P SDVFPVSLMWGYEY1LGRKOPAT1L1VDFMFN1TSKVA1TSEAS1PVE1K1TG1YKKE E8AH1L1KAP1Q1KQAD1FVSKFEND1W41D1GVSS1C1RGSE1T1QGD1HGD1RFRKFL EROP1LNDPDD1SSHYHNTSGSVAA1L1VPPAL1LSGF1Y1LK1K1R1K1P1K1YNGY AGHENSQCA5FENP1D1T1LK1TEK1AVR1D1T1N1CTV1W1"
misc_feature	484. .801
misc_feature	/note="Region: CUB domain 1"
misc_feature	823. .996
misc_feature	/note="Region: sushi domain 1"
misc_feature	1012. .1317
misc_feature	/note="Region: CUB domain 2"
misc_feature	1435. .1611
misc_feature	/note="Region: sushi domain 2"
misc_feature	1624. .1950
misc_feature	/note="Region: CUB domain 3"
misc_feature	1972. .2133
misc_feature	/note="Region: sushi domain 3"
misc_feature	2143. .2460
misc_feature	/note="Region: CUB domain 4"
misc_feature	2482. .2655
misc_feature	/note="Region: sushi domain 4"
misc_feature	2665. .2982
misc_feature	/note="Region: CUB domain 5"
misc_feature	3010. .3171
misc_feature	/note="Region: sushi domain 5"
misc_feature	3181. .3504
misc_feature	/note="Region: CUB domain 6"
misc_feature	3526. .3693
misc_feature	/note="Region: sushi domain 6"
misc_feature	3703. .4020
misc_feature	/note="Region: CUB domain 7"
misc_feature	4042. .4212
misc_feature	/note="Region: sushi domain 7"
misc_feature	4222. .4542
misc_feature	/note="Region: CUB domain 8"
misc_feature	4564. .4734
misc_feature	/note="Region: sushi domain 8"
misc_feature	4744. .5061
misc_feature	/note="Region: CUB domain 9"
misc_feature	5083. .5256
misc_feature	/note="Region: sushi domain 9"
misc_feature	5266. .5583
misc_feature	/note="Region: CUB domain 10"
misc_feature	5614. .5787
misc_feature	/note="Region: sushi domain 10"
misc_feature	5797. .6117
misc_feature	/note="Region: CUB domain 11"
misc_feature	6139. .6306
misc_feature	/note="Region: sushi domain 11"
misc_feature	6316. .6633
misc_feature	/note="Region: CUB domain 12"
misc_feature	6655. .6822
misc_feature	/note="Region: sushi domain 12"
misc_feature	6832. .7146
misc_feature	/note="Region: CUB domain 13"
misc_feature	7168. .7341
misc_feature	/note="Region: sushi domain 13"
misc_feature	7351. .7677
misc_feature	/note="Region: CUB domain 14"
misc_feature	7690. .7866
misc_feature	/note="Region: sushi domain 14"
misc_feature	7879. .8052

misc_feature	/note="Region: sushi domain 15" 8065. .8247
misc_feature	/note="Region: sushi domain 16" 8260. .8421
misc_feature	/note="Region: sushi domain 17" 8434. .8595
misc_feature	/note="Region: sushi domain 18" 8608. .8769
misc_feature	/note="Region: sushi domain 19" 8782. .8958
misc_feature	/note="Region: sushi domain 20" 8971. .9132
misc_feature	/note="Region: sushi domain 21" 9154. .9315
misc_feature	/note="Region: sushi domain 22" 9328. .9492
misc_feature	/note="Region: sushi domain 23" 9505. .9672
misc_feature	/note="Region: sushi domain 24" 9685. .9846
misc_feature	/note="Region: sushi domain 25" 9859. .10020
misc_feature	/note="Region: sushi domain 26" 10042. .10206
misc_feature	/note="Region: sushi domain 27" 10219. .10386
	/note="Region: sushi domain 28"

Alignment Scores:	
Pred. No.:	0
Score:	11353.50
Percent Similarity:	78.75%
Best Local Similarity:	63.78%
Query Match:	66.84%
DB:	9
	9
	63
	Gaps:
	9
	9
	1580
Length:	1580
Matches:	1981
Conservative:	465
Mismatches:	597
Indels:	63
Gaps:	9

US-10-016-248-2 (1-3104) X AF333704 (1-11580)

QY 1 Meaaiaaglyalaprofioalaaleuen-----LenProCyserien 15
 | | | | | |
 | | | | | |
DB 1180 ATCGTGCCACGGAACTCCATCATMTGGCTAAGTCGATGAACCTTCCCACTT 1239

10 -----cyscysalaserasnlnarghisser 27

[illegible][illegible]

QY

48 LySLeuMetProSerLysAspAsnSerG]nl.vgThrServa]l.euThrc]nv.a]c].v.a] 67

Db 1357 AAGATGCTGCCCAGCAGGATGGAAGCCATAAACTCTGTCTTGAGCCAGGAGGT 1416

68 SerGlnGlyHisAsnMetCysProAspProGlyIleProGluArgGlyLysArgLeuGly 87

141 / GCATGGTCTCATGTGTCTAGATCCTGGGATTCCAGAAATGGTAGAAGACAGGT 1476

[illegible][illegible][illegible]

127 HisArgProValCysArgAlaArgMetCysAspAlaHisGluArgGlyProSerGlyValLeu 146

Db 1597 CACAGGCCATCTGCCGAGGAGACATGTGATCCATCTGCGTGGGCCCAAGCGCGTC 1656

147 IIEThrSerProAsnPhenroileGlnTyraaspasnaAlahiscysValTyrIleile 166

Db 1657 ATTACCTCCCTAATTATCCGGTTCAGTATGAAATATGACACCTGTGTGGGTATC 1716
 Qy 167 ThrAlaLeuAnpProSerIyValIleValLeuAlaPheGluGluPheAspLeuGluArg 186
 Db 1717 ACCACCAACCAACCCGCAAGGTCATCAACCTTGTAAAGATTGAGTCGAGCGCA 1776
 Qy 187 G1YrAspThrLeuThrValG1YAspG1YG1YLeuAspG1YAspG1YLeuValLeu 206
 Db 1777 GGGTATACACCTCGACGCTTGGTATGCTCGGAGAGTGGAGACACCAATCGCTTGG 1836
 Qy 207 TyrMetSerGlnAsnAlaCysSerAspSerProHisThrProGlySerArgIleProGlu 226
 Db 1837 TACGTG-----CTCACGGATCCAGTCTTCCTGAC 1866
 Qy 227 -----SerMetSerGlyAspIleTyrArgGlnIlyThrPheValLeuGluIleCys 243
 Db 1867 CTGATTGTGACATGACCAACCAAGTGTGCTACAT----- 1902
 Qy 244 ArgAspIleSerSerAspAlaArgSerGlySerValArgIlySerProIlyThrSer 263
 Db 1903 -----CTGCAGTCGGATGATAG----- 1920
 Qy 264 AsnAlaValGluLeuValAlaProGly-----ThrGluIleGluGlnGly 278
 Db 1921 -----ATTGGCTCACCTGGGTTTAAAGCTGTTTACCAAGAAATTGAAGAGCA 1968
 Qy 279 SerCysGlyAspProGlyIleProAlaTyrGlyArgArgGluGlySerArgPheHis 298
 Db 1969 GGGGTGGGGATCTCGAATCCCGCTATGGGAGGAGCGGACAGCTTCCCTCAT 2028
 Qy 299 G1YAspThrLeuIyPheGluCysGlnProAlaPheGluLeuValG1YLeuAlaIle 318
 Db 2029 GGGATATCACTACCTTGAATGCCCGCGCTTTGAGCTGTGGGGAGAGATTATC 2088
 Qy 319 ThrCysGlnIyAsnAsnGlnIyPheSerAlaIyAspProGlyCysValPheSerCysPhe 338
 Db 2089 ACCGTGACGAGAACATCATGTGCTGCGCAACAGCCAGCTGTGATTTTCATGTTTC 2148
 Qy 339 PheAsnPheThrSerProSerGlyValValLeuSerProAsnIyProGluAspIyGly 358
 Db 2149 TTCAACTTACGGGATCATGTGGATTAATCTGTCAACCAATTAATCCAGGAAATAGCG 2208
 Qy 359 AsnHisIleuIyCysValIyPheIleuIleuAlaArgProIySerArgIleHisIleuAla 378
 Db 2209 AACACATGAACTGTCTGTGATTAATCTCGAGCCAGAACTGCATTCACCTTATC 2268
 Qy 379 PheAsnAspIleAspValGluProGlnPheAspPheLeuValIleIyAspGlyAlaThr 398
 Db 2269 TTTATATGATTTGATGTGAGCCCTCATTTGACTTCTCGCGGTCAAGATGATGCAAT 2328
 Qy 399 AlaGluAlaProValLeuGlyIyThrPheSerGlyAsnGlnLeuProSerSerIleThrSer 418
 Db 2329 TCTGACATTAACCTGCTCGGTAACCTTTCTGCGCAATAGTCCCTCCAGCTGGCCAGC 2388
 Qy 419 SerGlyHisValAlaArgLeuGluPheGlnThrAspHisSerThrGlyIyAspGlyPhe 438
 Db 2389 AGTGGGATATAGTCCCTTGAATTCAGTCTGACATTCACCTACTCGCAGAGGGTTC 2448
 Qy 439 AsnIleThrPheThrThrPheArgHisAsnGlnCysProAspProGlyValProValAsn 458
 Db 2449 AACATCACTTAACCAATTTGTCAGAAATGAGTGCATATCTGCAATTCCTATTAAC 2508
 Qy 459 G1YIyAspArgPheGlyAspSerLeuGlnLeuGlyIySerSerIleSerPheLeuCysAspGlu 478
 Db 2509 GAGACGAGCTTTGGTGAAGGTTTCTACTCGGAGCTCGGTTCTTTCACCTGATGAT 2568
 Qy 479 G1YPheLeuGlyIyThrGlnIySerGlyIyThrIleThrCysValLeuIyGluIySerVal 498
 Db 2569 GGGCTTTTCAAGACCGAGGATCCGAGTCCATTAACCTGCATACGCAAGAGGGAGTGG 2628
 Qy 499 ValIyPheAsnSerAlaValIleuArgCysGlyIyAlaProCysGlyIyHisIleuThrSerPro 518
 Db 2629 GTCTGAGCTCCACCGTGGCCCGCTGTGAAGCTCATGTGTGACATCTGACAGCGTCC 2688

Qy 519 SerGlyThrIleLeuSerProGlyIyTrpProGlyPheTyrIyAspAlaLeuSerCysAla 538
 Db 2689 AGCGAGTCAATTTGCTCTCGATGATGCCAGATATTAAGATTTTACTTATGAA 2748
 Qy 539 TrpValIleGluAlaGlnProGlyIyTrpIleIyHisIleThrPheAspArgPheIyThr 558
 Db 2749 TGGATTAATTTGAAGCAAAACCAAGCCACTATCAAAATGACTTTTACAGATTTACAGCA 2808
 Qy 559 GluValAsnTyrAspThrLeuGluValArgAspGlyArgThrTyrSerAlaProLeuIle 578
 Db 2809 GAGGTCAATTAATGACACTTGGAGTGCAGAGATGGCCAGCAAGTTCGCCACATGATC 2868
 Qy 579 GlyValTyrHisGlyIyThrGlnValProGlnPheLeuIleSerThrSerAsnIyLeuTyr 598
 Db 2869 GGGGATACCAAGGACCCAGGACCCAGTCTTCTATGACACCGGAACTTCAATGAC 2928
 Qy 599 LeuLeuPheSerThrAspIySerHisSerAspIleGlyPheGlnLeuArgIyGluThr 618
 Db 2929 CTGCTGTTCAACACTGACAAACCGCGCTCCAGATGGCTTCTCATCACTATGAGAT 2988
 Qy 619 IleThrLeuGlnSerAspHisCysLeuAspProGlyIleProValAsnGlyIyAlaHis 638
 Db 2989 GTGACCCCTTGAGTGGATTCCTGCTGAGACCCGGGATCCCTGTGAACGGCCATCCCCAC 3048
 Qy 639 G1YAsnAspPheThrValG1YAlaLeuValIyThrPheSerCysAspSerGlyIyThrLeu 658
 Db 3049 GGTGAGACTTGGCAATCAAGTCCACAGTATCTTAAGCTGTGACCCGGGATCACACTA 3108
 Qy 659 SerAspGlyIyGluProLeuGluCysGluProAsnPheGlnIyTrpSerArgAlaLeuProSer 678
 Db 3109 AGTGAAGACAGAGCCCTCGTCTGTGAGAGAAACACAGATGAAACACACCGCTTGGCCAG 3168
 Qy 679 CysGlnAlaLeuCysGlyIyPheIleGlnIySerSerGlyIyThrIleLeuSerProGly 698
 Db 3169 TGGCAGCTCTATGTGAGAGCTTACATCAAGGAGAAAGTGAACAGTCTTCTCTCGGG 3228
 Qy 699 PheProAspPheThrProAsnAsnLeuAsnCysThrIleIleGluThrSerHisGly 718
 Db 3229 TTTCCAGATTTTATCAAACTCTTAACCTGACAGTGAACATTAAGTCTTCAATGGG 3288
 Qy 719 IyGlyIyValPhePheThrPheHisIyThrPheHisIleGluIySerGlyIyHisAspIyLeuLeu 738
 Db 3289 AAAGAGTCAATGATATCTTCAACACTTTCATCTTGAAGTTCACACGACTATTTACG 3348
 Qy 739 IleThrGluAsnIySerPheThrGlnProLeuArgGlnLeuThrIyGlySerArgLeuPro 758
 Db 3349 ATCACAGAGATGAGGATTTTTCGAGCCCGTTCAGAGCTCACCGGGTCTGGTGTGCT 3408
 Qy 759 AlaProIleSerAlaGlyLeuTyrGlyAsnPheThrAlaGluValArgPheIleSerAsp 778
 Db 3409 CATACATCAAGGACAGCCCTGTTGAAACTTCACTGACCCAGCTTGGTTTATATAGAC 3468
 Qy 779 PheSerMetSerIyGluGlyPheAsnIleThrPheSerGlyIyTrpAspLeuGluProCys 798
 Db 3469 TTTCTCAATTTGTCAGAGGCTTCAATATCACTTTTCAAGAAATGACTGAGGCATGT 3528
 Qy 799 GluGluProGluValProAlaTyrSerIleArgIyGlyLeuGlnPheGlyValGlyAsp 818
 Db 3529 GATGATCTGAGAGTCCCTGCTTCAGCCGAAGAAATGGTTTTCATTTGGTGTGGAGAC 3588
 Qy 819 ThrLeuThrPheSerCysPheProGlyIyTrpArgLeuGluGlyIyThrAlaArgIleThrCys 838
 Db 3589 TCTCTAGCTTTTCTCGCTTCTCGGATATTCGTTTGAAGAGTCCACCAAGCTTACTCTG 3648
 Qy 839 LeuGlyIyArgArgArgLeuTrpSerSerProLeuProArgCysValAlaGluCysGly 858
 Db 3649 CTGGTGGGGGGCGCGCTGTGTGAGTGCACCTCTGCCAAGTGTGTGGCCAAATGGA 3708
 Qy 859 AsnSerValIyThrGlyIyThrLeuLeuSerProAsnPheProValAsnTyrAsn 878
 Db 3709 GCAAGTGTCAAGGAATGAAAGAAATTAATCTGTCCAAATTTTTCATTCATTAATGAT 3768

Db	4849	GACCTTGGCACCGCTTGATATTTCAAAAGTTTCAACATGGAGCCACGCTATGACTTCTTA	490
Qy	1259	HisIleTyrAspGlyAaAspSerLeuSerProIleuIleGlySerPheTyrGlySerGln	127
Db	4909	CACATCTATGAAGGGGAAGTATCCAAACAGCCCTCATTTGGAGATTACAGGGCTTCAG	490
Qy	1279	LeuProGlyArgIleGlyIleuSerSerSerAsnSerLeuPheIleuAlaPheArgSerAspAla	123
Db	4969	GCCCAAGAAAGATATGAGAGTAGACCGAAACAGCTGTTCTGGCATTTGGAGAGAGCC	502
Qy	1299	SerValSerAsnAlaGlyPheValIleAspTyrThrGluAsnProArgGlySerCysPhe	131
Db	5029	TCCCTGGCGCTTTCACAGGCTTCGCATGTGAATTAAAGAAACCAAGGAAAGCTGTTTT	508
Qy	1319	AspProGlySerIleLeuAsnGlyTyrAsnGlyValGlySerAspLeuValLeuGlySerSer	133
Db	5089	GACCGAAGAAATATATGAATATGGACACAGAGTTGGAAACACACTTCAGCTTCACCC	514
Qy	1339	ValThrTyrTyrCysHisGlyGlyTyrGluValGluGlyThrSerThrLeuSerCysIle	135
Db	5149	ATCACCTTACAGATGTGACTCTGGCTATTAAGATTCTTGACCCCTCATCCACACTGTGTG	520
Qy	1359	LeuGlyProAspGlyIysProValITrpAsnAsnProArgProValCysThrAlaProCys	137
Db	5209	ATTGGGGCTGATGGGAAACCTCTCGAGCAAGTCTGCCCTCCGCAATGCTCCCTGT	526
Qy	1379	GlyGlyGlnTyrValGlySerAspGlyValValLeuSerProAsnTyrProGlnAsnTyr	139
Db	5269	GGAGGCAAGTACACGGGATCAGAGGGGATGTTTATACACAACTACCCCAATATTAC	532
Qy	1399	ThrSerGlyGlnIleCysLeuTyrPheValIThrValProIysAspTyrValValPheGly	141
Db	5329	ACAGCTGGTCAATATATGCTCTATTCATCCACAGGTACCAAGAAATCTGTGGCTTTGA	538
Qy	1419	GlnPheAlaPhePheHisThrAlaLeuAsnAspValValGluValHisAspGlyHisSer	143
Db	5389	CAGTTTGCTATTTCCAGACACCCCTGATGATTTTGGCAAAATTATTATGGAACCCAT	544
Qy	1439	GlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThrGlyGlyIleuSerProIleu	145
Db	5449	GCAACAGCCACGACTTCTACGCTACCTCGGGGTCTCACTCAGGGGAAACATTGCCCTTG	550
Qy	1459	AlaThrSerAsnGlnValLeuTyrLeuPheSerAlaValGlyLeuAlaProAlaArgGly	147
Db	5509	GCTAGCTCAATCAATCAATTCTGCTCGATTCACTGCAAGAGGGGTGCTCTCCCGGCG	556
Qy	1479	PheHisPheValTyrGlnAlaValProArgThrSerAlaThrGlnCysSerSerValPro	149
Db	5569	TTCCACTCTGTGTATCAAGCTGTTCTCTGTACAGTGAACCCATATGACGCTCTGCC	562
Qy	1499	GluProArgTyrGlyValArgLeuGlySerAspPheSerValGlyAlaIleValArgPhe	151
Db	5629	GAGCCCAATACGGAAGAGAAATTGTTCTGAGTTTCTGCCGGTCCATGCTCGATTTC	568
Qy	1519	GlyCysAsnSerGlyTyrAlaLeuGlnGlySerProGluIleGlyCysLeuProValPro	153
Db	5689	GAGTGCACACCGGGGATACCTGCTTCAAGGTTTCCACGGGCTCCACTGCGACGCGTGCC	574
Qy	1539	GlyAlaLeuAlaGlnITrpAsnValSerAlaProThrCysValValProCysGlyValAsn	155
Db	5749	AACGCTTGGCACAGTGAAGACACACGATCCCAAGCTGTGTGGTACCTTGACGGCAAT	580
Qy	1559	LeuThrGluArgArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSer	157
Db	5809	TTCACTCAACAGAAAGGTACAAATCTCTGCCCCGGGTACCTTGAGCCATACGAAACAC	586
Qy	1579	LeuAsnCysValITrpyValIleValValProGluGlyAlaGlyIleGlnIleGlnValVal	159
Db	5869	TTGAACGTATATGAAAGATCAATATGTTACGAGGGCTCGGAATTTCAAGATCCAAGTAT	592
Qy	1598	IleSerPheValIThrGlnGlnAsnITrpAsnSerLeuGluValPheAspGlyAlaAspAsn	161

Db 5929 CAGTTTTCGACGAGGAGAACTGGGACTTCCTTTGAGATCCAGATGGTGGGAGATGT 5988
 Qy 1618 ThrValThrLeuGlySerPheSerGlyThrThrValProAlaLeuAsnSerThr 1637
 Db 5989 ACCGACCCAGACTGGGAAAGCTTCTCAGGACCAAGTACCGGACGCTGAACAGTACT 6048
 Qy 1638 SerAsnGlnLeuValPheThrSerAspIleSerValSerAlaIleGlyPheHis 1657
 Db 6049 TCCAACCACTCTACCTGCATTTCCAGTGTACATTTAGTGGGAGAGCTGCTGGTTCCAC 6108
 Qy 1658 LeuGlnTyrValThrValGlyLeuSerSerCybProGlnProAlaValProSerAsnGly 1677
 Db 6109 CTGGAATAACAAAGTGTAGGTCTGTGCATGCAAGCAACAGAGCCCTCCAGCAAGAC 6168
 Qy 1678 ValValThrGlyGlyValGlyTyrLeuValAsnAspValValSerPheGlnGlyValProGly 1697
 Db 6169 ATCAAAATCGGAGATCGGTATAGTGTAAACAGATGCTCTCTCCAGTGGAGACCCGG 6228
 Qy 1698 TyrAlaLeuGlnGlyHisAlaHisIleSerCybMetProGlyThrValArgArgTyrAsn 1717
 Db 6229 TACACCTTCGACGGCCGCTCCACATTTCTGTATGCGAGGACCGTTCCGCTTGGAC 6288
 Qy 1718 TyrProProProLeuCybIleAlaGlnCybGlyThrValGlnGlnMetGlyVal 1737
 Db 6289 TATCCGTCTCCCTGTGCATTCGCAACCTGTGAGGAGCGCTGACACCTTGGGTGTGTG 6348
 Qy 1738 IleLeuSerProGlyPheProGlyAsnTyrProSerMetMetArgProSerThrValIle 1757
 Db 6349 ATCTGTAGCCCGGCTTCCAGGTCTTACCCCAACAACTTGAAGCTGAGAGATC 6408
 Qy 1758 AlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGlnProAsn 1777
 Db 6409 TCATTAACCACTGGCTATGTGTCATTTCAATTTCTGAAATTTTCTACCGAAGCTAT 6468
 Qy 1778 HisAspTyrIleGlnIleArgAsnGlyProGlyGlnThrSerArgMetMetGlyArgPhe 1797
 Db 6469 CATGACTCTCTTGAATTCAAATGGAACCTTACCAACAGCCCATGATGGACAATTT 6528
 Qy 1798 SerGlySerGlnLeuProSerSerLeuLeuSerThrSerHisGlnThrThrValTyrPhe 1817
 Db 6529 AGCGGACGAGATCTCCCGCGGCTCTGTGAGCAACAGCATGAACCTTCATCAGCTT 6588
 Qy 1818 HisSerAspHisSerGlnAsnArgProGlyPheLeuSerGlnGlyThrAlaTyrGlnLeu 1837
 Db 6589 TATGTGACCATTTGCGAAACCGGACAGAGATTTAACTGCTTACCAAGCTATGATTA 6648
 Qy 1838 GlnGlnCybProAspProGlnProPheAlaAsnGlyIleValArgGlyAlaGlyTyrAsn 1857
 Db 6649 CAGAACTGTCCAGATCCACCCCATTTCCAGAAATGGGTATCATGCACTCGGATTAAGC 6708
 Qy 1858 ValGlnGlnSerValThrPheGlnCybLeuProGlyTyrGlnLeuThrGlyHisProVal 1877
 Db 6709 GTGGGCAATCAGATCTTTGAGTGTATCTGGGTATCATTTAAATGAGGCATCTGTC 6768
 Qy 1878 LeuThrCysGlnHisGlyThrAsnArgAsnTyrAspHisProLeuProGlyCysGlnVal 1897
 Db 6769 CTCCTTGTACAGCTGGGATCAACAGAACTGGAATCACTTTTCCAAAGATGATGCC 6828
 Qy 1898 ProCysGlyGlyAsnIleThrSerSerAsnGlyThrValTyrSerProGlyPheProSer 1917
 Db 6829 CTTGTGGGTACAGTAACTTCTCAGAAAGGACCATCTACTCCCTGGCTTCCGAT 6888
 Qy 1918 ProTyrSerSerSerGlnAspCybValThrPheLeuIleThrValProIleGlyHisGlyVal 1937
 Db 6889 GAGATATCGATCTGAAAGACTGATTTGGCTCATCAGGTGCTCCAGGGCAGAGATT 6948
 Qy 1938 ArgLeuAsnLeuSerLeuGlnThrGlnProSerGlyAspPheIleThrIleTyrAsp 1957
 Db 6949 TACATCAACTTCACTCTTACAGAGGAACTGTCAACATTTCACTTGTCTTGGGAC 7008
 Qy 1958 GlyProGlnGlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaValGlyThr 1977
 Db 7009 GGTCCGATCAGAACTCAACCCAGCTGGGAGTTTTCAGTGGCAACAGCCCTCGAAACG 7068

Qy 1978 ValGlnSerSerSerAsnGlnValLeuLeuValPheHisArgAspAlaIleThrGlyGly 1997
 Db 7069 GCGTATAGCTCCCAACCAAGTCTGCTCAAGTTCCACAGCACTTTTCAAAATGAGAGC 7128
 Qy 1998 IlePheAlaIleAlaPheSerAlaTyrProLeuThrLysCybProProProThrIleLeu 2017
 Db 7129 TTTCTTGTCTCAATTTCCACGCAATTTCAAGTCAAGAAATGTCACTCCCGCAGCGGT 7188
 Qy 2018 ProAsnAlaGlnValValThrGlnGlnGlnPheAsnIleGlyAspIleValArgTyr 2037
 Db 7189 CCAAGGCAAAATGCTTACTGAGATGATGAAATTTGAAATAGAGATTTTGTGAAGTAC 7248
 Qy 2038 ArgCysLeuProGlyPheThrLeuValGlyAsnGlnIleLeuThrCysLysLeuGlyThr 2057
 Db 7249 CAGTCCACCCCGGATACCTTGTGGGAGCCAGCATTTGTGACTTGCAGAGCTCAGTTC 7308
 Qy 2058 TyrLeuGlnPheGlnGlyProProProIleCysGlnValHisCybProThrAsnGlnLeu 2077
 Db 7309 CAGTTGAGTTTGAAGGTTCTCTCCCAATGTGAAGCAACATGCCAGCAAAATGAAGTC 7368
 Qy 2078 LeuThrAspSerThrGlyValIleLeuSerGlnSerTyrProGlySerTyrProGlnPhe 2097
 Db 7369 CGGACTGATCATCGGAGCATTCATCCAGTCAAGGATCCGGGTATTTTAACTCC 7428
 Qy 2098 GlnThrCysSerThrLeuValArgValGlnProAspTyrAsnIleSerLeuThrValGln 2117
 Db 7429 CAGACTTGTCTTGAAGTATTAAGTGAACCAACTACAACTTACATCTTTGTGGAC 7488
 Qy 2118 TyrPheLeuSerGlnGlyGlnTyrAspGlnPheGlnIlePheAspGlyProSerGlyGln 2137
 Db 7489 ACATTTCAAGTGAAGAAAGCACTTTGATGCACTGGAGTGTATGTTGTTCTTGGGCA 7548
 Qy 2138 SerProLeuLeuValAlaLeuSerGlyAsnTyrSerAlaProLeuIleValThrSerSer 2157
 Db 7549 AGTCTCTGTAGTACTTAAAGTGGAAACATCACTGACATCAATTTTACAGAGC 7608
 Qy 2158 SerAsnSerValTyrLeuArgTyrPheSerSerAspHisAlaTyrAsnArgLysGlyPheLys 2177
 Db 7609 AGTATAGTTATATCTCCCTGTGTCTGACGACCAAGTCCACAGTAAAGAAATTCAG 7668
 Qy 2178 IleArgTyrSerAlaProTyrCysSerLeuProArgAlaProLeuHisGlyPheIleLeu 2197
 Db 7669 ATTGCTATCAGACACTTCTGCAATTTGACCAACCCCTGAAGATGGGGATTTCTA 7728
 Qy 2198 GlyGlnThrSerThrGlnProGlyGlySerIleHisPheGlyCysAsnAlaGlyTyrArg 2217
 Db 7729 AACAGGACTGCAGAGCGGTTGGAGCAAGTGCATTAATTTTTCAGACCTCGATACCGA 7788
 Qy 2218 LeuValGlyHisSerMetAlaIleCysThrArgHisProGlnGlyTyrHisLeuTyrPhe 2237
 Db 7789 ATGGTGGCCCAAGCAATGCAACCTGTGAAGCAAAACCACTTGGCATGTACAGTGGAGC 7848
 Qy 2238 GlnAlaIleProLeuCybGlnAlaLeuSerCysGlyLeuProGlnAlaProLysAsnGly 2257
 Db 7849 TCCCTCAGGCACTCTGCCAGGCTGTCTGTGGATATCCAGATATCCAGGAACGCT 7908
 Qy 2258 MetValPheGlyTyrGlyTyrThrValGlyThrLysAlaValTyrSerCybSerGlyGly 2277
 Db 7909 TCATTTTCCGGAACGAGTTCACCTTGGACAGTAAATGCTCATGATGATCTGAGAGC 7968
 Qy 2278 TyrHisLeuGlnAlaGlyAlaGlnAlaThrAlaGlnCysLeuAspThrGlyLeuTyrPhe 2297
 Db 7969 TTCAAGTTAAATCCAGCCCAAGCAACAGCCGTGTGTCAAGAAATGGGCTGTGGAGT 8028
 Qy 2298 AsnArgAsnValProProGlnCysValProValThrCysProAspValSerSerIleSer 2317
 Db 8029 AACAGGGAAGAGCGCCCATATGTGAAGCGGTGCTTGGCCCAAGCATGAAGCTCAGCTC 8088
 Qy 2318 ValGlnIleGlyArgTyrPheGlyLeuIlePheGlnThrGlnTyrGlnPheGlnAlaGlnLeu 2337
 Db 8089 TCAGAACATGTCAATCGAGAGCTGTGTTTCAGAGATCTTGAATGATGATGAGTGTCTCAAGTA 8148

QY 2338 MetLeuIleCyAspProGlyTyrTyrThrGlyGlnArgValIleArgCysGluAla 2357
Db 8149 TTGGAGAGCTGACGCTGGTTACTTAAAGGCTGGAGGCTCTCGGTCGACAGGCC 8208
QY 2358 AsnGlyLysTrpSerLeuGlyAspSerThrProThrCysArgIleIleSerCysGlyGlu 2377
Db 8209 AATGGAGAGTGAACATGAGAGATGAGAGGCCAAGCTGTCAATCTCGTGGGAAGC 8268
QY 2378 LeuProIleProProAsnGlyHisArgIleGlyThrLeuSerValTyrGlyAlaThrAla 2397
Db 8269 CTTTCCTCTCCCAATGGCAACAGATGGAACTGGACATTTATGGGGCCACAGCT 8328
QY 2398 IlePheSerCysAsnSerGlyTyrThrLeuValGlySerArgValArgGlyCysMetAla 2417
Db 8329 ATATTACGTGCACACCGGCTACACGCTTGTGGGCTCATGTCAAGAGTGGTTGGCA 8388
QY 2418 AsnGlyLeuTrpSerGlySerGlyValArgCysLeuAlaGlyHisCysGlyThrProGlu 2437
Db 8389 AATGGGCTGGAGACGGGACGGAACCTGATGTCTGGCTGGCCACTCGGTTCCCGACAG 8448
QY 2438 ProIleValAsnGlyHisIleAsnGlyGluAsnTyrSerTyrArgGlySerValIleTyr 2457
Db 8449 CCGATTGGAGACGGTCACTTAAGTGAAGATGGCTTCACTTACAGACACAGGTTTAC 8508
QY 2458 GlnCysAsnAlaGlyPheArgLeuIleGlyMetSerValArgIleCysGlnAsnPhs 2477
Db 8509 CAGTGCATCTCGGTTCCGGCTTGTGGAACTTCGAGAGATATCGCTGCAAGACAC 8568
QY 2478 HistPheSerGlyLysTrpProPheCysValProIleThrCysGlyHisProGlyAsnPro 2497
Db 8569 AATGGTCTGGACAAACGCTGTCTGTCTCCCACTCACTAGTGTCACTCGTGAACCCCT 8628
QY 2498 ValAsnGlyLeuThrGlnGlyAsnGlnPheAsnLeuAsnAspValValLysPheValCys 2517
Db 8629 GCCACCGATTCATTAATGAGAGTGAAGTCACTGATATGTCTGGAATTTCACTCGC 8688
QY 2518 AsnProGlyTyrMetAlaGlyGlyAlaAlaArgSerGlnCysLeuAlaSerGlyGlnTrp 2537
Db 8689 AACACGGGCTAATTTGCTGACGGGGGTCTCTGAGCCAGTGTGCGAACAACGGCCAGTGTG 8748
QY 2538 SerAspMetLeuProThrCysArgIleIleAsnCysThrAspProGlyHisGlnGluAsn 2557
Db 8749 AGTAGCCCTCTGCCACAGTGTCAAGTGTGAACGTGTGATCAAGGCTTTGTGGAAAT 8808
QY 2558 SerValArgGlnValHisAlaSerGlyProHisArgPheSerPheGlyThrThrValSer 2577
Db 8809 GCCATTGCTGACGGGCAACAGAACTTCCCTGAGAGTTTGAGATGAATGATGATCCG 8868
QY 2578 TyrArgCysAsnHisGlyPheTyrLeuLeuGlyThrProValLeuSerCysGlnGlyAsp 2597
Db 8869 TACCATTTGCAAGAAAGGATTTCACTGTGGATCTTCAAGCCCTGTATGGCAAAAT 8928
QY 2598 GlyThrTrpAspArgProArgProGlnCysLeuLeuValSerCysGlyHisProGlySer 2617
Db 8929 GGGTTATGGAGACCGATCCCTGCCAAGTGTGGCTATATCTGTGGAACCCAGGGGCTC 8988
QY 2618 ProProHisSerGlnMetSerGlyAspSerTyrThrValGlyAlaValAlaArgTyrSer 2637
Db 8989 CCGCCAACGCGCTCTCACTGAGAGAGCTGTTAATCTATGCGCGCTGACACTACCTCC 9048
QY 2638 CysIleGlyLysArgThrLeuValGlyAsnSerThrArgMetCysGlyLeuAspGlyHis 2657
Db 9049 TGCAGAGGAGAGACGACCTCATAGGACACAGACAGAGAGTGGCAGGAAACAGCTAC 9108
QY 2658 TrpThrGlySerLeuProHisCysSerGlyThrSerValGlyValCysGlyAspProGly 2677
Db 9109 TGGAGCGGGGCACTGCCCACTGCAACAGAAATTAATCTGATCTGTGTATATCCGGGG 9168
QY 2678 IleProAlaHisGlyLysArgLeuGlyAspSerPheAspProGlyThrValMetArgPhe 2697
Db 9169 ACCCCAGCACTGGGCTCGGCTGTGTGATGATTTAAGACAAAGAGCTTCTCCGGCTTC 9228
QY 2698 SerCysGlyAlaGlyHisIleValLeuArgGlySerSerGlyValArgThrCysGlnAlaAsnGly 2717

Db 9229 TCTTGAAATAGGGGACCACTGAGGGGCTCCCTGAACGACGTTTGTCTCAATGGG 9288
QY 2718 SerTrpSerGlySerGlnProGluCysGlyValIleSerCysGlyAsnProGlyThrPro 2737
Db 9289 TCATGGTCAAGACTGAGCGGCTGTGTAGAGCGCGTCTCTGTGGCAACCTTGACACACC 9348
QY 2738 SerAsnAlaArgValAlaPheSerAspGlyLeuValPheSerSerIleValTyrGlu 2757
Db 9349 ACCAAGGAATATGTCAAGTGTGATGATGATCTTCTTCCAGCTCGGTCACTATGACC 9408
QY 2758 CysArgGlnGlyTyrTyrAlaThrGlyLeuLeuSerArgHisCysSerValAsnGlyThr 2777
Db 9409 TCGTGGAAAGGCTTACAAAGCTCAGGGCTCATGACACGCAATTCACAGCCATGGAGACC 9468
QY 2778 TrpThrGlySerAspProGluCysLeuValIleAsnGlyAspProGlyIleProAla 2797
Db 9469 TGGACAGGCACTGCCCGACTGACAAATTAATTTATGTTGGGATCCAGGCACTAGCA 9528
QY 2798 AsnGlyLeuArgLeuGlyAsnAspPheArgTyrAsnLysThrValThrTyrGlnCysVal 2817
Db 9529 AATGGATCAAGTTGGACCGACTTCACTTCAACAAAGACTGAGCTATCACTAGTAAAC 9588
QY 2818 ProGlyTyrMetMetGlySerHisArgValSerValLeuSerCysThrLysAspArgThr 2837
Db 9589 CCAAGCTATGTCAATGGAAGCACTCAATCCGCACTATTCGCTGTATCCAAAGACGGCAGG 9648
QY 2838 TrpAsnGlyThrLysProValCysValAlaLeuMetCysLysProProProLeuIlePro 2857
Db 9649 TGGAAATCCGAGCAACCTGTCTGCAAAACCGTGTCTGTCTCTGACGCCCGCGGTGAG 9708
QY 2858 AsnGlyLysValValGlySerAspPheMetTrpGlySerSerValThrTyrAlaCysLeu 2877
Db 9709 AATGGAACAGTGAAGGAAGTATTTCCGCTGGGGCTCAGCAATAGTTAAGCTGCATG 9768
QY 2878 GlnGlyTyrGlnLeuSerLeuProAlaValPheThrCysGlnGlyAsnGlySerTrpThr 2897
Db 9769 GACGGTTACAGACTCTCACTCCGCCATCTCTCTGTGAAGTCCGGGGGTGTGGAA 9828
QY 2898 GlyLeuLeuProGlnCysPheProValPheCysGlyAspProGlyValProSerArgGly 2917
Db 9829 GAGAGATCCCGCAAGTGTCTCTGTGTCTGCGAGAGCCCTGAGCATCCCGCGAAGAGG 9888
QY 2918 ArgArgGluAspArgGlyPheSerTyrArgSerSerValSerPheSerCysHisProPro 2937
Db 9889 CGACTTATGGGAAAGTTTACCTTAATGATGACCAAGTCTTCTTCCAGTGCATAATCTCCA 9948
QY 2938 LeuValLeuValGlySerProArgArgPheCysGlnSerAspGlyThrTrpSerGlyThr 2957
Db 9949 TTTATACTCGTGGATCTCCAGAAAGTGTGCCAAGTGCAGGCACTGGACGCGGATTA 10008
QY 2958 GlnProSerCysIleAspProThrLeuThrThrCysAlaAspProGlyValProGlnPhe 2977
Db 10009 CAACCACTGCAATGATCTGCTCATTAACAACCTGCGCAAGCCCTGGATACGCCACACTT 10068
QY 2978 GlyIleGlnAsnAsnSerGlnGlyTyrGlnValGlySerThrValIlePheAspCysGln 2997
Db 10069 GGAATACAGAAATTAAGTCTCAAGAGGTTAGAGTTGGAAGACGTTTTCAGGTGACAGA 10128
QY 2998 LysGlyTyrLeuLeuGlnGlySerThrThrArgThrCysLeuProAsnLeuThrTrpSer 3017
Db 10129 AAAGGCTACATATTCAAGGTTCCAGACACTGCGACCTGTCCTGCAATTAACATGAGT 10188
QY 3018 GlyThrProProAspCysValProHisHisCysArgGlnProGluThrProThrHisAla 3037
Db 10189 GGGATACAGACCGCAATATATCTCATGCTGCGAGAGACGCAAGAACCCGGCAGACGGG 10248
QY 3038 AsnValGlyAlaLeuAspLeuProSerMetGlyTyrThrLeuIle-ThrProAlaArgAr 3057
Db 10249 GATGTGAGAGGACCATTTCTTCCACTTGGGCTACACCTTAAGTACACTGCGACATCA 10308
QY 3057 GAlaSerProSerArgValAlaProSerThrAlaProAlaArgArgMetAlaAlaGlyG 3077

Db 10309 GCGTTTCTCCGAGGGGATCTGACACAGACATGTAAAGCAGACATGAATGACA 10368
Qy 3077 palaserArg 3080
Db 10369 GGAAGTCC 10378
RESULT 5
AY017475
LOCUS AY017475 11221 bp mRNA linear ROD 17-JUL-2001
DEFINITION Mus musculus CsmD1 (CsmD1) mRNA, complete cds.
ACCESSION AY017475
VERSION AY017475.1 GI:14787175
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Sun, P. C., Uppaluri, R., Schmidt, A. P., Pashia, M. E., Quant, E. C.,
1 (bases 1 to 11221)
Sun, P. C., Uppaluri, R., Schmidt, A. P., Pashia, M. E., Quant, E. C.,
2 (bases 1 to 11221)
TITLE Transcription map of the 8p23 putative tumor suppressor region
JOURNAL Genomics 75 (1-3), 17-25 (2001)
MEDLINE 21365705
PUBMED 11472063
REFERENCE
AUTHORS Sun, P. C., Uppaluri, R., Schmidt, A. P., Davis, M. E., Quant, E. C.,
Sunwoo, J. B., Gollin, S. M., and Scholnick, S. B.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-2001) Ocular oncology, Washington University
School of Medicine, Box 8115, 517 S. Euclid Ave, Saint Louis, MO
63110, USA
FEATURES
source location/Qualifiers
1..11221
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
1..11221
/gene="CsmD1"
309..11003
/gene="CsmD1"
/note="similar to Homo sapiens CUB and Sushi multiple
domain 1 (CsmD1)"
/codon_start=1
/product="CsmD1"
/protein_id="AAG54083.1"
/db_xref="GI:14787176"
/translation="MTAMRKFSLLPLVLAVLCAGLLTAAGQCCGGLVQNGRTIE
SPGPHGYPNYANCTWIIITGERRIQLSTFTALREDDPILAVDGPQGNLKYRL
SGRLPSSTYSTGSLTLMTPTDPAVAQAQKMYEVLPSHTGCPBEILLKYLHGR
FNIDKIRYSCLSGYLEGHAITLCIVSPGASWDPAFCRAEGACGGLRGTSGS
ISSEPFSEYDNNADCTWTLAEPDITAVFDPLVEGDEPLEISGTAPSWLWG
MNLPSPISSKNNRLHFTSDSNRRKAPNAQFVKALIELKSGVALLSKSSHN
SVLQGVSLSDMCPDGPJPDNGRRAGSDPRVAVNFQSCEDYVLAQKGLTCOV
TETLAANDHPRICRARTCGSNLRGSPGIVTSPPYGVSDYLNACWVLTTPDDKVI
KLAEFEELRGVYTLTVGDAKGVGRSVLTVSGSVDPDLVSKSNQMLHQSND
SIGSGFKAVQETIEKGCQDGPJPAVGKRTGSSFLHGDITLTCOAFLVGRVIT
CQKNQMSGNKPSQVSCFNFETAPSGIILSPNPEYGNMNCVMILISEPGRHL
IPKDPVEPODFLAVKDDGISDITVIGTSGNVPAQLASSGHIVLLEFOSDSTG
RGFNITVTTFGQNCNDPGLPVNGRGDRPLGSSVSPHCDDFVKTQSESLTL
QDGNVWSTVPRCEAPCGGLTSSGVLTPGMPGVYKDSINCEWIEAKPGHSIT
TDFRFQTEVAVDTIEYRDGPTSSPLIGEHGTAPPLISTGYMLTLTTDSRSRS
VGLIHYESVTLSDSLDPLGIPVNGRHGNSFIRSTVFPSCPGYTLDDDELVCE
KNHQMNHALPSCDALCGGVIHGKGTVALSPDPYVNSINCTVIEVSGKVQMP
HTFHLSSHDLTLTEDGSFEPVPAVLPTIIRKAGLFGNFTQALRFTISPSISY
EGFNITFAEYDLRCPDGPVAFRRIGFQGVGDTLAFCTCFQLEGATKLTLGAG
GRVWASAPLRPCVAGSAYKNGEITLSPVPSYNNHNECIYKITEAGKGIHARA
RTQLRFGDTLKYVDKQSSRSIGVTRSRPMYLVANSNTIRLEFNTNGSTPAG
FQLYTIFDLVKCEPGLPNTYGRIRIDRGHTDVLVLSCPGAMHGSSITLCLSD
RRVWDKMPSCVACCGGLVHAATSGRLSPYAPYNNHACTVITIEADGKTIISHF
IVPDETAHDLIKWMDGVDSNILLKEMSGALPEDIHSTFNSLTLOPDSDFPISKSG
FSIQFSTIASIENDPGMPONGTRGYSREGDITITFCQDGVQLQPAKITCYOLNN
RFWQDPDPSICIAACGNLTGSPAVIILSPNYPQYVPPGKCDMIRIKNPPVIALIRK

SFMEPSYDFLIHYEGEDSNPLIGSFQGSQABERIESSGNSLFLAFRSDASYLGSF
AIEFKRPRACDFPDGNINNGRITGTDPLKSGVTVQCDGKYIVPDSIECTYQGA
KPSMDRLPACQAPCGGOYVTSRGVLSLSPYPNHAYAGCMQVSIYPERKVEGDA
YFOTALLDLABELDGTTHPOARLSSLSGSHSGTTLPLATSNQILLRPSAKSGASARF
HFVYQAVPRTSDTQSCSSVPEPRYGRRIIGSFSGASIVREPCNCGVILLOSTAIRCOSV
PNALQOMNDTIPSCVPCSGNFTORRGITILSPGYPEPYNNLNCVMKIIVSEBSGLOI
OVISFATQEWMDLEIHDGDMTAPRLGSGFTGTVALLNSTNSQCLHFOSDISVVA
AGHLRYKTVGLAACQEPALPSMGIKIGDGYVNDVILSFCEGSGYTLQGSNLSHSCMG
TYRRNMYPSPLCATCGGTITSMSGVILSGFPGSGPNNIDCTMKISLIPGCAHIOF
LNFSTANDHYLEITONGPYHSSPMWQDSFQSLISTLTHETTLRFRSDHSONROG
FKLSYQAYELONCPDPAFQNGFMINDSVSGSISFECPGVILLGLPVLTCQHDV
RMNMYPEPRCDACGVNVSQNGTIVSPGPDPYLLKQCLMLVTPPGHGVINFTL
LQTEAVNDVYAVWDGPDONSPOLGVSNGTAPETAYSTNOVILKHSDFNSGPFVL
NFHAPOLKPCPPPAPOADLTLEDDEFDGDKVYKQCHGPTLLLSDLTCLLSQL
LFQGSPTCEACQCPANBYRESSGVILSPYPNYNSQTCMSIYKRPNTITLPLD
TFQSEKQFDLAEVFDGSSGRSPLVLVLSGHTESQNSFTSKSNHLYRMWSDTHSKKG
FKIRYAPYCSLSTLRNGOILNKRTAGAVSKVHYCKPQYRMIGHSNATCRANPVGV
YQWDSMAPLCOAVSCGIPKAPNGSGFTGNEFTLDSKYVYECNCGFLDASQEAITYQ
EDLWMSNRGKPTCKPVPSPISGOLSEHVMRLVSGSLNEGYAOVLSCSPYFLOG
ORLLCOQANGTSTEDRPRCKYISGSGISFPNGNKGITLITVYATAIPTCNQYTL
VGSHTVECLANGMSSETRCLAGHCGSPYINGHISDGFERYPTVYQCNPRPL
VGSVRICCRITSGRRLIVCPITCGHPENPAHGLNGTEFNLDIWNFTCHGRIL
QGSARAQCRNSQWSSPLICRVVNCSDPSVSNAYARHQGNPSEFEVGSYWHCK
TGFYLLGSSALTCMASGLMDRSLPKCLAISCGHPYANAVLTELFTYCATVQYCK
GGOILTGNSTRVCOEDSHMSGLPHCSGNSPGCGPDPYPAHSGRLGDEKTKSLRF
SCMGHOLRGFAERTCLVNGSMGVOVPCVAGSAGNPGPTNMLISDGLPSSAYI
YACMEGKITSGLMTRHCTANGTGTAPDCTIISCDPGLPLNGIOPGDPFTNKYS
YQCNPGILMEPTSPITRCKDGTNNQSRPLCAVLCNODPPVNGVBSDBRWGMS
ISYSCVDGYLSSALISCGRGVWGEVQCLPVCGDGPABEGLSKSTFSE
VFIQCKPEFVLVSSRRTCQADWISGIIPTCIDPAHTACPDGTHFGIONSKEYE
VGSITVPRCKGHIOGSTRCLANLTWMSGIOTECIPACRQEPETPAHADVAIDLP
AGVTLVYTCHPGPFPLAGSEHRTCKADMKWMSKYSCKYREVNETVTPVPD
VFFINSVMKYIYELGRQDPTITVDFNATSKNANPAAARVOLIELGYKKEBA
HLLKAPHLKGPADIVFSKENDNMGLDVSGLRGGSSFGQDHGKRGFKLER
QDPSNDADSSNNHYOGTSSGVAAILVPEFALLISGFAYILVHRRTPRYOYNGAG
HENSQOASFEENMYDTNLKPTBAKAVREFDTINTVCTVV"
402..719
/gene="CsmD1"
/note="Region: CUB domain 1"
741..914
/gene="CsmD1"
/note="Region: sushi domain 1"
930..1235
/gene="CsmD1"
/note="Region: CUB domain 2"
1353..1526
/gene="CsmD1"
/note="Region: sushi domain 2"
1539..1665
/gene="CsmD1"
/note="Region: CUB domain 3"
1687..2048
/gene="CsmD1"
/note="Region: sushi domain 3"
2058..2375
/gene="CsmD1"
/note="Region: CUB domain 4"
2397..2570
/gene="CsmD1"
/note="Region: sushi domain 4"
2580..2697
/gene="CsmD1"
/note="Region: CUB domain 5"
2925..3086
/gene="CsmD1"
/note="Region: sushi domain 5"
3096..3419
/gene="CsmD1"
/note="Region: CUB domain 6"
3441..3608
/gene="CsmD1"
/note="Region: sushi domain 6"
3618..3935

US-10-016-248-2 (1-3104) X AY017475 (1-11221)

QY	1	MeAlaIGlyAlaProProProAlaLeuLeu-----LeuProCysSerLeu	15
Db	1098	ATAGGGGAGCAGAAAGCAACCATTCATATGGCTCAGACGATGAACCTCCATCCCTGTGG	1157
QY	16	11SerAsp-----CysCybAlaSerAsnGlnArgHisSer	27
Db	1158	ATCAGCAGCAGAAAGACTGGGTCCGTCTCCACTTTACATCTGACAGCAATCACCAGCTAA	1217
QY	28	ValGlyValGlyProSerGluLeuValIleYslySglnIleGluLeuYsSerArgGlyVal	47
Db	1218	---GGGTTCATAGCTCAGTTCAGATGCAATGTGAAAAAGCAATGTGAGTCAAGTCCAGAGGAGTC	1274
QY	48	LysLeuMetProSerLysAspAsnSerGlnYsThrSerValLeuThrGlnValGlyVal	67
Db	1275	AGATGCTTACCAAGACAGACGACGACCAAAAACCTTGCTTGACCCAAAGAGGGGTG	1334
QY	68	SerGlnGlyHisAsnMetCysProAspProGlyIleProGluArgGlyLysArgLeuGly	87
Db	1335	TCTTCATCTCAGACATGTCAGATCCGGAGTTCCAGCAATGGCAGAAAGCAGAGCT	1394
QY	88	SerAspPheArgLeuGlySerSerValGlnPheThrCysAsnGlnGlyTyrrAspLeuGln	107
Db	1395	TCGATTTGAGGGTGGTGCAAAATGTGMAATTCCTGTGAGCAATTCGTAATCTC	1454
QY	108	GlySerLysArgIleThrCysMetLysValSerAspMetPheAlaIleTPSerAspHis	127
Db	1455	GGTGCCAAAGGAAATCAGCTGTCAAAAGATCAGGAGACCTTGCTGATGGAATGATC	1514
QY	128	ArgProValCysValArgAlaArgMetCysAspAlaHisLeuArgGlyProSerGlyIleIle	147
Db	1515	AAACCCATCTGTGCGACCGACACATGTGGATTCMAAACCTGGTGGCCAACTGGTGTCATC	1574
QY	148	ThrSerProAsnPheProIleGlnTyrrAspAsnAsnAlaHisCysValTPPileIleThr	167
Db	1575	ACCTCCCTTAATCTACCCAGTGCATATGAAGACATCTCACTGTGTGGATCATTACA	1634
QY	166	AlaLeuAsnProSerLysValIleLysLeuAlaPheGlnIubPheAspLeuGluArgGly	187
Db	1635	ACAACTGATTCACGACCAAGTCAATCAACCTTGTAAGAGTTTGAGTTGGAAACGAGG	1694
QY	188	TyrAspThrLeuThrValGlyAspArgGlyGlnAspGlyAspGlnYsThrValLeuTyrr	207
Db	1695	TAACGACCCCTGACAGTGGGAGATCGTGGGAGATGGGGGACCAACGATCCGCTTGAT	1754
QY	208	MetSerGlnAsnAlaCysSerAspSerProHisThrProGlySerArgIleProGlu---	226
Db	1755	GTG-----CTCACAGGCTCCAGTGTCTCTGACCTTC	1784
QY	227	---SerMetSerGlyAspIleThrArgGlnIleYsThrThrValLeuGlnIleCysArg	244
Db	1785	ATCGTAGCAATGACCAATCAGATGTGGCTTCAC-----	1817
QY	245	AspIleSerSerSerAspAlaArgSerGlySerValArgLysSerProLysThrSerAsn	264
Db	1818	---CTGCAGTCCAGTGCACGC-----	1835
QY	265	AlaValGluLeuValAlaProGly-----ThrGluIleGlnGlnLysSer	279
Db	1836	-----ATTGGTACACAGAGGTTTAAAGCTGTACCAAGAAATTGAGAAAGGAGGC	1886
QY	280	CysGlyAspProGlyLleProAlaTyrrGlyArgArgGlnGlySerArgPheHisHisGly	299
Db	1887	TGTGGGACCTTGCGATCCAGCCTTATGGGAGCGGAGCTGGCAGCAGCTTTTGATGGA	1946
QY	300	AspThrLeuLysPheGlnCysGlnProAlaPheGluLeuValGlyGlnIleValIleThr	319
Db	1947	GACACGCTTACCTTTGAGTCCAGCAGAGCGTTTGAAGCTGTGGAGAGAGACTATTACC	2006
QY	320	CysGlnLysAsnAsnGlnTPSerAlaLysLysProGlyCysValPheSerCysPhePhe	339
Db	2007	TGCCAGAAAACACACAGTGGTGGGCAACAAGCCAAAGCTGTGTTCATGATTTCTTC	2066
QY	340	AsnPheThrSerProSerGlyValValLeuSerProAsnTyrrProGluAspTyrrGlyValn	359

```

Db 2067 AACCTCAGGACCCCTCGGATCATCTCTCGCCAACTATCTGTAGAGATAGGCAAC 2126
Qy 360 HileuHieCyVal1TTPLeu1IleuAlaArgProGluSerArg1Leu1IleuAlaPhe 379
Db 2127 AACATGAACCTGTGTGTGTATATATCTGAGCTCGGAGCCGAGATTCACCTCATCTTC 2186
Qy 380 AsnAspIleAspVal1GluProGlnPheAspPheLeuVal1IleuAspArg1Val1ThrAla 399
Db 2187 AAGGATTTTCAGTGTGAAGAACTCAGTTGACTTCTCGCCGCAAGATGATGGATTTCT 2246
Qy 400 GluAlaProVal1LeuGly1ThrPheSerGly1AsnGlnLeuProSerSer1LeuThrSerSer 419
Db 2247 GACATCACAGTCCCTCGAACTTCTCTGCAATAGAGTCCCTGCACAGCTGCCAGCAGT 2306
Qy 420 GlyH1sVal1AlaArgLeuGlnPheGln1ThrAspH1SerThrGly1ValArgGly1PheAsn 439
Db 2307 GGACACATAGTACCCCTCGAGTTTCAGTGTGATCATCTCCACCTGGCAGAGTTTCAAC 2366
Qy 440 IleThrPheThrThrPheAspH1sAsnGlnCyAspProAspProGly1Val1ProVal1AsnGly 459
Db 2367 ATCATATATACCAATTTGTCAGAACAGTGCACAGCCTGGAGATTCCTGTAAATGGA 2426
Qy 460 LysArgPheGly1AspSerLeuGlnLeuGly1SerSer1LeuPheLeuCyAspGluGly 479
Db 2427 CGGCGTTTGGAGACAGGTTTCTGCTGGAGATTCAGTCTCTTCCAGTGTGATGATGCC 2486
Qy 480 PheLeuGly1ThrGlnGly1SerGly1Thr1LeuThrCyVal1LeuVal1Gly1SerVal1 499
Db 2487 TTTGTGAAGACTCAGGTTCTGAGTCCATCATCTGATCTTGAAGATGAAATGTGGTTC 2546
Qy 500 TrpAsnSerAlaVal1LeuAspCyGluAlaProCyGly1Gly1IleuThrSerProSer 519
Db 2547 TGGAGCTTACTGCTCTCGCTGTGAAGCTCATGTGTGGGCACTGACGCGTTCCAGT 2606
Qy 520 Gly1Thr1LeuSerProGly1TTPProGly1PheThrLysAspAlaLeuSerCyAla1TTP 539
Db 2607 GGGGTATATTAACCTCCAGAGTGGCCGGGATACATAAAGATTCCTTAAATGTGATGATGG 2666
Qy 540 Val1IleGluAlaGlnProGly1TTPPro1Leu1IleuThrPheAspArgPheLeuThrGlu 559
Db 2667 GTCATTTGAGCAAAACAGACGATTCCTCAAGATTAACATTCGACAGGTTCCAAACAGAA 2726
Qy 560 ValAsnThrAspThrLeuGlnVal1ArgAspGly1ArgThr1TyrSerAlaProLeu1IleGly 579
Db 2727 GTCAATTATGATACCTCGAAGTACGGGATGGCCAAACACCTGATCCCACTATGAGA 2786
Qy 580 Val1TyrH1sGly1ThrGlnVal1ProGlnPheLeu1IleSerThrSerAsn1TyrLeuThrLeu 599
Db 2787 GAGTACCAAGGACCCAGGACCAACAGTTCCTCATACAGACAGGCACTACATGATCTG 2846
Qy 600 LeuPheSerThrAspLysSerH1sSerAsp1IleGly1PheGlnLeuAspGly1GluThr1Le 619
Db 2847 CTGTTTACACTGATAGACAGCCGGGCTGAGCTTGCTTCCTCACTATGAGAGTGTG 2906
Qy 620 Thr1LeuGlnSerAspH1sCyLeuLeuAspProGly1LeuProVal1AsnGly1GlnArgH1sGly 639
Db 2907 ACTTTGAATCCGACTCTGTCTGACCCAGGACATCCCTGTAAATGTGACAGCCGATAGGC 2966
Qy 640 AsnAspPheThrVal1Gly1AlaLeuVal1ThrPheSerCyAspSerGly1TyrThrLeuSer 659
Db 2967 AGTACCTTTGGTATCAGGTCTACAGTACCTTACAGCTGTGTGACCTGGGTACACACTCAGT 3026
Qy 660 AspArg1GluProLeuGlnCyGly1ProAspPheGln1TTPSerArg1AlaLeuProSerCyS 679
Db 3027 GACGATGAGCCCTCGTCTGTGAGAAACATCATGAGTGAACACAGCTTCCAGAGCTGT 3086
Qy 680 GluAlaLeuCyGly1Gly1Phe1IleGlnGly1SerSerGly1Thr1LeuSerProGly1Phe 699
Db 3087 GATCTCTGTGTGAGGCTACATCATGGAAGAGTGGAGTGTCTTCTCCAGGATTT 3146
Qy 700 ProAspPheThrProAsnAsnLeuAsnCyThr1TTP1Ile1Gln1ThrSerH1sGly1Lys 719

```

```

Db 3147 CCAGACTTTTATCCAACTCTGAACTGATGACATCGACATCGAAGTCTCATGCGAAG 3206
Qy 720 GlyVal1PhePheThrPheH1sThrPheH1sLeuGlnSerGly1AspThr1LeuLeu1Le 739
Db 3207 GAGGTGAGATGAAATTTCCACACCTTTCACCTTGAAGATTTCCACGACTATTTGTATTC 3266
Qy 740 ThrGluAsnGly1SerPheThrGln1ProLeuArg1GlnLeuThrGly1SerArg1ProAla 759
Db 3267 ACAGAGATGAGGATTTCTCAGAGCCAGTACCGACAGCTCAGTGGTGTAGTCTGCTCAC 3326
Qy 760 Pro1LeuSerAlaGlyLeuThrGly1AspPheThrAlaGlnVal1ArgPhe1IleSerAspPhe 779
Db 3327 ACCATTAGAGCTGGTTGTTTGGAATCTTACATCGACACATCGACAGTTCATCTGATTC 3386
Qy 780 SerMetSer1TyrGlnGly1PheAsn1IleThrPheSerGly1TyrAspLeuGlnProCyGlu 799
Db 3387 TCCATCTCCATAGAAAGCTTCAACATTCATGTCGCAAAATATGACCTGGAGCCCTGTGAT 3446
Qy 800 GluProGluVal1ProAla1TyrSer1LeuArgLysGly1LeuGlnPheGly1Val1AspThr 819
Db 3447 GACCTTGAAGTCCCTGCTTCCAGCCGCAAGATGGTTTCAGTTCCGTTGGTGGTGAACAC 3506
Qy 820 LeuThrPheSerCyAspPheProGly1TyrArgLeuGlnGly1ThrAlaArg1LeuThrCyLeu 839
Db 3507 CTGGCTTTCACTGCTTCCAGGATACCGCTTAGAAGTGCACCAAGCTTACCTGCTGCTG 3566
Qy 840 Gly1Gly1ArgArg1GlnLeu1TTPSerSerProLeuProArgCyVal1AlaGly1CyGly1Asn 859
Db 3567 GGTGGGGAGCCGAGGTGAGTGCACCTGTGCAAGGTGTGTGCTGATGTGAGCA 3626
Qy 860 SerVal1ThrGly1ThrGlnGly1ThrLeuLeuSerProAspPheProVal1Asn1TyrAsn 879
Db 3627 AGCTCAAGAAATGAATGAAGAACTTCTCTTCCAAATTTCCATGCCATATGATAT 3686
Qy 880 AsnH1sGly1Cys1Ile1TyrSer1IleGln1ThrGlnProGly1LysGly1IleGlnLeuVal1 899
Db 3687 AACCATAGATGTATCTATAAATAGAAACGAAGCCGGAAGGGGATCCATCTCAGAGCT 3746
Qy 900 ArgAlaPheGlnLeuSerGlnGly1AspVal1LeuLysVal1TyrAspGly1AsnAsnAsnSer 919
Db 3747 CGAACCTTCCAGCTTTCGAAGGAGACACTTAAAGGTTTACATGGAAGAAAGACAGTCC 3806
Qy 920 AlaArgLeuLeuGly1Val1PheSerH1sSerGly1MetMetGly1Val1ThrLeuAsnSerThr 939
Db 3807 TCAAGTCACTGGGTGCTTTCACAGAGCGGAATTCATGGGGCTGGTCTCACAGACACC 3866
Qy 940 SerSerSerLeuThrPheAspPhe1IleThrAspAlaGluAsnThrSer1Gly1PheGlu 959
Db 3867 TCCAACTACCTGAGGCTGAGTTCAACACACAGCGGTCCGACACCGCCCAAGCTTCAG 3926
Qy 960 LeuH1sPheSerSerPheGlnLeu1LeuCyGly1AspProGly1TTPProLysPheGly 979
Db 3927 CTCACCTACACCGATTTTGACTAGTGAATGTGAGATCCAGGCACTCTTACATAGGC 3986
Qy 980 Tyr1LysVal1H1sAspGlnGly1H1sPheAlaGly1SerSerVal1SerPheSerCyAspPro 999
Db 3987 TACAGATCCGTGTAGTATGGCACTTACACAGACCTTATAGTTCTCTACAGTGTCAACCG 4046
Qy 1000 Gly1TyrSerLeuAspArg1Gly1SerGlnGluLeuLeuCyLeuSerGly1GluArgArgThr1TTP 1019
Db 4047 GGCATACCATGATGACAGCAGATGACCTTACCTGATGAGTGGAGACCGGAGGTGTGG 4106
Qy 1020 AspArgProLeuProThrCyVal1AlaGly1CyGly1Gly1ThrVal1ArgGly1GluVal1Ser 1039
Db 4107 GACAAACCTATGCTTCTGTGTGGCTGGAATGTGTGTGTGTGCTTGTGTCATGACGACCATCG 4166
Qy 1040 Gly1GlnVal1LeuSerProGly1TyrProAlaPro1TyrGlnH1sAsnLeuAsnCy1Leu1TTP 1059
Db 4167 GGGCGCATGCTGTCTGTGTGTACCGCCCATATACAAACCTTCTACCTGACCTGG 4226
Qy 1060 Thr1IleGluAlaGluAlaGly1CyThr1IleGly1LeuH1sPheLeuVal1PheAspThrGlu 1079
Db 4227 ACCATTAAGGCGCATCTGTGAAGAACATGACGCTCATTTATGTGTGTGACACTGAA 4286

```

```
QY 1080 GIuValHieAspValIleuArgIleThrAspGlyProValGluSerGlyValIleuLeuLys 1099
Db 4287 ACGGGCGCATGACATCTCAAGGTCTGGAGTGGCCCTGTGAGACAGACATCTGCTGTGAG 4346
QY 1100 GIuLeuSerGlyProAlaLeuProIleuAspLeuHisSerThrPheAsnSerValIleu 1119
Db 4347 GAGTGGAGCGGTTCGGCCCTCACTGAGGACATCCACACCTTCACTCACTCACTG 4406
QY 1120 GluPheSerThrAspPhePheThrSerLysGlnGlyPheAlaIleGlnPheSerValSer 1139
Db 4407 CAGTTTGACAGTACTTCTTCATCCAGCAAGTCCGGCTTCTCCATCCAGTTTCCACTTCC 4466
QY 1140 ThrAlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerGlyAspSer 1159
Db 4467 ATTCGATCCACCTCGCAATGACCTCGGATGCTCCAGATGCTACCCGTCAGTGAACG 4526
QY 1160 TrpGluAlaGlyAspSerThrValPheGlnCysAspProGlyIleValAlaLeuGlnGlySer 1179
Db 4527 CGGGAACCTGGAGACACATCACTCCAGTGTACCTGGATACAGCTCCAGCTCCAGGACA 4586
QY 1180 AlaGluIleSerCysValIleValIleGluAsnAspPhePheThrGlnProSerProThr 1199
Db 4587 GCCAAGATCAGCTGTGTCACTGACCTGAACACCGTTTCTGTGGCAGCAGACCTCCCTTCG 4646
QY 1200 CysIleAlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeuSerProAsn 1219
Db 4647 TGCATAGCTGCTGTGTGGGATCTGACAGCCCTGCTGGAGATTCATCCCAAC 4706
QY 1220 TyrProGluProIleProProGlyIleGlyCysAspTrpIleValThrValSerProAsp 1239
Db 4707 TACCACAGCCATACCCCTCTGGAGAGAGTGCATCGAGAGATTAAGTGAACCCAGC 4766
QY 1240 TyrValIleAlaLeuValPheAsnIlePheAsnLeuGluProGlyIleAspPheLeuHis 1259
Db 4767 TTGTGATGCTGTGATTAATTAAGTTTAAAGTGAAGTTCAGGAGCTCTCAAGCC 4826
QY 1260 IleTyrAspGlyAspAspSerLeuSerProLeuIleGlySerPheTyrGlySerGlnLeu 1279
Db 4827 ATCTATATAGGGAGGAGACTCCACAGCCACTGATTGGAAGTTTCCAGGAGCTCTCAAGCC 4886
QY 1280 ProGlyIleArgIleGluSerSerSerAsnSerLeuPheLeuAlaPheArgSerAspAlaSer 1299
Db 4887 CCAGAGAGGATCCAGAGACAGTGGTAACAGCTCTTCTGTCATTCAGAGTGAATGCTTCT 4946
QY 1300 ValSerAsnAlaGlyPheValIleAspTyrThrGluAsnProArgGluSerCysPheAsp 1319
Db 4947 GTTGGCTGTGGGATTCGCCATTTGAATTAAAGAAACCCCGGAGAGCTTGTGAC 5006
QY 1320 ProGlySerIleLysAsnGlyThrArgValGlySerAspLeuLysLeuGlySerSerVal 1339
Db 5007 CCTGGGAACATTAATGAATGGGACAGAGATTGGAACCTTAAACTGGAGCTCTCAAGAT 5066
QY 1340 ThrTyrTyrCysHisGlyGlyIleGlyValGluGlyIleThrSerThrLeuSerCysIleLeu 1359
Db 5067 ACCTATCAGTGCATCTGCTTACAGATTGTGATCCCTGCTCCATTTAGTGTGTGACA 5126
QY 1360 GlyProAspGlyLysSerProValTrpAsnAsnProArgProValCysThrAlaProCysGly 1379
Db 5127 GGGGCTGATGGGAGGCGCTCTGCGGACCGGCGACTACCTGCGCCAGAACACCTGTTGA 5186
QY 1380 GlyIleTyrValGlySerAspGlyValIleLeuSerProAsnTyrProGlnAsnTyrThr 1399
Db 5187 GGGCATATACCGGGCTCGAGGGGGTGTGTTGTGTCACCAACCTCACTATTAATCAACA 5246
QY 1400 SerGlyGlnIleCysLeuTyrPheValThrValProLysAspTyrValValPheGlyGln 1419
Db 5247 GCTGGGCAAGTGTGCTTATTCATCAACGATGCGCAAGAAATGTGTGTATTTGACAG 5306
QY 1420 PheAlaPhePheHisThrAlaLeuAsnAspValIleGluValHisAspGlyHisSerGln 1439
Db 5307 TTTCCTATTTCCAGTCACTGACAGCACTTGAGAGAGTGTGTGATGGAACACACACCC 5366

QY 1440 HisSerArgLeuLeuSerSerLeuSerGlySerHisThrGlyGluSerLeuProLeuAla 1459
Db 5367 CAGCCAGAGCTTCTCACTTCTCTGCTGCTCCATTCAGGTGAACACATGCGACAGCTGCT 5426
QY 1460 ThrSerAsnGlnValLeuIleLysPheSerAlaIleGlyLeuAlaProAlaArgGlyPhe 1479
Db 5427 ACATCCAAATCAATTTGCTGTGATTCAGCGCCAAAGGGAGCTTCTGACGGGCTTTC 5486
QY 1480 HisPheValTyrGlnAlaValProArgThrSerAlaThrGlnCysSerSerValProGln 1499
Db 5487 CACTTGTCTACCAAGCTGTTCAGGACCAATGATACATGACATGACCTGTCTCCCGAG 5546
QY 1500 ProArgTyrGlyLysArgLeuGlySerAspPheSerValGlyAlaIleValArgPheGln 1519
Db 5547 CCCAGATATGGAGAGAGATTTGTTCTGATTTCTCGAGCTCCATCGTCCGCTTCGAG 5606
QY 1520 CysAsnSerGlyTyrAlaLeuGlnGlySerProGlnIleGlyCysLeuProAlaProGly 1539
Db 5607 TGTAAACCAAGTTTACCTGCTGCAAGTTTCCACCGCATCCGCTGCACTGTGCAAT 5666
QY 1540 AlaLeuAlaGlnTrpAsnValSerAlaProThrCysValAlaProCysGlyLysAsnLeu 1559
Db 5667 GCTTTGCTCAGTGAATGACACATCCCAAGCTGTGATTCATGACAGTGGCACTTC 5726
QY 1560 ThrGluArgArgGlyThrIleLeuSerProGlyPheProGluProTyrLeuAsnSerLeu 1579
Db 5727 ACTCAGAGAGAGGAGCAATTTGTCCTCAGCATATCTGAGCCCTATGGAGAACAACTG 5786
QY 1580 AsnCyValTrpLysIleValAlaProGluGlyAlaGlyIleGlnIleGlnValValSer 1599
Db 5787 AACTGTGTGGAAGATCATGTATCGAGGGCTCAGGATCCAGATTCAGATCAATCAAGTCAC 5846
QY 1600 PheValThrGluGlnAsnTrpAspSerLeuGluValPheAspGlyAlaAspAsnThrVal 1619
Db 5847 TTGGCACAAGACAGAACTGGAGCTCCCTGAGATCCATGATGAGAGACATGACGAGC 5906
QY 1620 ThrMetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerThrAsn 1639
Db 5907 CCGAAGCTGGGAGGCTTCTCAGTACCAAGATGCTGACCTGCAACAGCACTCCAAAC 5966
QY 1640 GlnLeuTyrLeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHisLeuGln 1659
Db 5967 CAGCTGCTGCACTTCATCCAGTCAATCATGATGCTGTGCTGCGCGCTTCACTGAG 6026
QY 1660 TyrIleThrValGlyLeuSerSerCysProGluProAlaValProSerAsnGlyValLys 1679
Db 6027 TACAAACGGTGGGCTGGCTGCGTGCAGAGAACCTGCTCCGACCAATGGCATCAG 6086
QY 1680 ThrGlyIleArgTyrLeuValAsnAspValValSerPheGlnCysGluProGlyTyrAla 1699
Db 6087 ATAGGAGACCGGTACATGATGTAACGATGTGTCTTCCAGTGTGAACCTGGGTATACC 6146
QY 1700 LeuGlnGlyHisAlaHisIleSerCysMetProGlyIleValArgArgTyrPheTyrPro 1719
Db 6147 TTGAGGGCGCTACACATTTCTGTATGCCAGGAACCGTACAGCTGCTGAACTACCT 6206
QY 1720 ProProLeuCysIleAlaGlnCysGlyGlyIleValGluLeuGluGlyValIleLeu 1739
Db 6207 TCCCTTTGTGATGCTGCACTGCTGTGGGACATCAACGCAATGAGTGAATGATCTG 6266
QY 1740 SerProGlyPheProGlyAsnTyrProSerAsnMetAspCysSerTrpLysIleAlaLeu 1759
Db 6267 AGCCCAAGGTTTCCAGGCTCATACCCCAACACTTGAATGACACCTGGAAATATCACTG 6326
QY 1760 ProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGluProAsnHisAsp 1779
Db 6327 CCAATTTGCTATGCTGCAATATCAATTTCTAAATTTCTCAACTGAAGCAACATGAC 6386
QY 1780 TyrIleGlnIleArgAsnGlyProTyrGluThrSerArgMetMetGlyArgPheSerGly 1799
Db 6387 TACCTGAGATTCAGATGAGCTTCAACAGAGTCCAAATGATGGCCAGTTCAGTGGC 6446
QY 1800 SerGluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTyrPheHisSer 1819
```


OY	2539	pWetLeuPProThrCysArgIleIleAsnCyserThrAspProGlyYhiSgIngluAenserva	2559
Db	8666	CCCCCTTGGCCATCTGTCCAGTGTGTAACGTTCCAGCCCGGATCTGTGGAAAAATGCGT	8725
OY	2559	lArgInValhIbalaSergIyProhIaArgPheSerPheGlyThrThrValSerTyar	2579
Db	8726	TCGGCATGGGCAACAGACTTCCAGAGAGTTTCCAGATATGGACAACTGTGATGATNCA	8785
OY	2579	gCyAsnhIbGlyPheTyrlleuGlyThrProValleuSerCySgIngluAenservIth	2599
Db	8786	CTGCAGAAACGGGGTCTACTGCTGGAGTCTTCCCTGACCTGCATGGCAGTGGCTT	8845
OY	2599	rTrpAspArgProArgProGlnCysLeuLeuValSerCySgIyhiAspProGlySerPropr	2619
Db	8846	GTGGGACCGCTCCTTACCAAGTGTCTGGCATATCATGTGGGACCTGGAGTCCCTGC	8905
OY	2619	ohIserGlnMetSergIyAspSerTyrlhValGlyAlaValAlaArgTyserCySgI	2639
Db	8906	TAATGCTCTCCTGACTGGAGAAATTGTTACATACGAGCCACGGTTCAGTACTCTGCA	8965
OY	2639	eGlyLysArgThrIleuValGlyAsnSerThrArgMetCySgIyLeuAspGlyhiSTpTh	2659
Db	8966	AGGGGGCCAGATCTTCACAGGCAACAGCAAGAGTGTGCCAGAGACAGTCACTGGAG	9025
OY	2659	rGlySerLeuPProhIaCyserGlyThrSerValGlyValCySgIyAspProGlyIlePr	2679
Db	9026	TGATATCCCTTCTCATTTGTTCAAGAAATAGTCCGATTTTGTGTATTCAGGAGACCC	9085
OY	2679	oAlaIhIbGlyIleArgleuGlyAspSerPheAspProGlyThrValMetArgPheSerCy	2699
Db	9086	AGCAATAGGGTCTGTGCTTGGGAGTGAATTAAACAAGAGTCTTTGGCTTCTCTCTG	9145
OY	2699	sGlyAlaGlyhIaValleuArgGlySerSergIuArgThrCySgInlalaSnhIySerT	2719
Db	9146	TGAATAGGCGCACAGCTGGGGCTTTCAGAACGACATGCTGGTAATAGATCTCG	9205
OY	2719	pSergIySergInProGluCysGlyValIleSerCySgIyAsnProGlyIthrProserAs	2739
Db	9206	GTCAAGAGTACAGCTGTGTGTGAGCGCTGCTCGGAGAACCCAGGACCGCGACCA	9265
OY	2739	nAlaArgValIaPheSerArgIyLeuValPheSerSerIleValTyrlGluCysAr	2759
Db	9266	CGGATATCTCTGACGATATGGATCTTCTTCAGTCTGTGATATATCTCCGCTG	9325
OY	2759	gGluGlyTyrlhValaThrGlyLeuLeuSerArgIhIaCyserValAsnGlyThrThrPh	2779
Db	9326	GGAAGGCTACAGACGTGGGGCTCAGACTCGGCACTGACGCGCCAAATGGAGCGTGC	9385
OY	2779	rGlySerAspProGluCysLeuValIleAsnCySgIyAspProGlyIthrProAlaAsnG	2799
Db	9386	AGGCACAGCTCTGACGACATCATCAGCTGTGGGATCTCTGGACACTGCCCAATGG	9445
OY	2799	yLeuArgLeuGlyAsnAspPheArgTyrlhValIthrValIthrTyrlGluCysValProG	2819
Db	9446	CATCCAGTTTGGGACAGACTTCTCTTCAACAAGACTGTGATATCAATGCAACCCCG	9505
OY	2819	yTyrlMetMetGluSerhIaArgValSerValleuSerCySgIyAspArgThrTrpAs	2839
Db	9506	CTACTAATAGAGCCCCCAACATCACCCATCCGCTGCACCAAGATGGCATGGAA	9565
OY	2839	ngIyThrIysProValCyAlaValAlaLeuMetCySgIySerpProProleuIleProAsnG	2859
Db	9566	TCAGACCGGCGCTCTGCAGAACTGTCTGTGCACACGCTCTCCAGTCCCAATATGG	9625
OY	2859	yLysValIaGlySerAspPheMetTrpGlySerSerValIthrTyrlAlaCyAlaGluG	2879
Db	9626	AAAGGTGAGGGGTCACTTCGATGGGGGTGCGCATTAAGCTACAGTGTGTGAATGG	9685
OY	2879	yTyrlGlnLeuSerIleuProAlaValPheThrCySgIyGlyAsnGlySerTTrpThGlyG	2899
Db	9686	CTACCAAGCTCTCCCACTCGGCATCTGTCTGTGCGAAGGCGTGAAGTGTGAAGAGGA	9745

OY	2899	uLeuPProGlnCysPheProValPheCySgIyAspProGlyValIProSerArgGlyArArg	2919
Db	9746	AGTCCCTCAGTGGCTTGGCTGTGTGTGTGTATCTGTGACATCCGCGCAGAGGAGCGCT	9805
OY	2919	gGluAspArgGlyPheSerTyrlhArgSerValSerPheSerCySgIhIaProProleuVa	2939
Db	9806	CAGTGGAAAAGCTTACCTTTAAGTCTAGGTCTTCACTCCAGTGCAGACCCCATTCGT	9865
OY	2939	IleuValGlySerProArgArgPheCySgInserAspGlyThrTrpSergIyThrGlnPr	2959
Db	9866	ATTAGTGGATCTCTCAGAGAAACCTGCCAGCGCATGGGATATGAGTGGCATCCAGCC	9925
OY	2959	oSerCySgIleAspProThrLeuThrThrCyAlaAspProGlyValIProGlnPheGlyI	2979
Db	9926	CAGTGTATAGATCCAGCACACACCGCTGCCCAACCTGTGCACCTCCCATTTGAAAT	9985
OY	2979	eGlnAsnAsnSergInGlyTyrlGlnValGlySerThrValleuPheArgCySgInlGly	2999
Db	9986	CCAGAAATAGCTTAAGGATATGAGTGTGGAGCACTGTGTTTTCAGATGTGAAGAAAG	10045
OY	2999	yTyrlleuGlnGlySerThrThrArgThrCySgLeuProAsnLeuThrTrpSergIyTh	3019
Db	10046	GTAACATATCCAAAGCTCCACGACGCGACCTGCTGCCAACCCTCACTGAGTGGAT	10105
OY	3019	rProProAspCySgValProhIaSerhIaCyArgGlnProGluThrProThrIbAlaAsnVa	3039
Db	10106	CCAGACGAGTGCATACCCCATGCTTGGCGGACAGCAAGACCCAGCGCATGACAGAT	10165
OY	3039	IglValaLeuAspLeuProSerMetGlyTyrlhThrLeuIleThrProAlaArgAlaAs	3059
Db	10166	GAGACCATTCAGTCTTCCAGCTTTTGGCTTACACTTATGCTACACTGCCACCTCGATT	10225
OY	3059	ePProSerArgValaIaProSerThrAlaProAlaArgArgMetAlaIaGlyAlaAs	3079
Db	10226	TTTCTTCTGCTGTGATCTGAACACAGGACGTGTAAGCAGATATGAATGACAGGAA	10285
OY	3079	er 3079	
Db	10286	GT 10287	

RESULT 6
 AX374896
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 XREFS
 KEYWORDS
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 TITLE
 JOURNAL
 Amgen Inc. (US)
 FEATURES
 source
 location/Qualifiers
 1..10433
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 <1..9303
 /note="unassigned protein product"
 /codon_start=1
 /protein_id="CAU26764.1"
 /db_xref="GI:19169791"
 /translation="TLTVGDAGKVGDTSRVLVLTGSSVPLDITVSNOMWMLHOSDD
 SISPGRKAYOEIERGGGDDPGIPAYGRGTSSPIHGNLTPECPALAEIVGERVIT
 COONMOSGNKBSCVFSCPFNPTASGILLSPYEPREXNNKNCWLIISEPSRHL
 IRNDPVEPQFDLAKDQISDITITLFTGFSNVEVSQALSSGHIVLRLEQSDHSTTG
 RGNITITTFGQNECHDPIGIPNGRGRFPLGSSVSHFCDDGFPVKTGSSITCIL
 QDNVAVSVTPRCAPCGHVLASGVILPDMPGYVDSDLSHCEWIIIAKPHSITKI
 TFDPRQTEVNYDTLAEVRDGPASSSPILGEYHGTQAPQLISTGNFMYLLFTDNRSS


```
QY 622 GlnSerAspHisCysLeuAspProGlyIleProValaAnglyGlnArgHisGlyAsnAsp 641
DB 1213 GAGTCGGATTCCTGCTGACCCGGGCACTCTTGAGAACGCCATCGCCACGGTGGAGAC 1272
QY 642 PheTyValGlyAlaLeuValThrPheSerCysAspSerGlyTyThrThLeuSerAspGly 661
DB 1273 TTGGGATCAGGTCCACAGTACTTTCAGCTGTGACCCGGGGTACACCTAAGTACACAC 1332
QY 662 GluProLeuGluCysGluProAsnPheGlnThrPheSerArgAlaLeuProSerCysGluAla 681
DB 1333 GAGCCCTCGTCTGAGAGGAGAACCAACAGTGAACCAAGCCCTTGCCAGCTGGAGACCT 1392
QY 682 LeuCyGlyGlyPheIleGlnGlySerSerGlyThrIleLeuSerProGlyPheProAsp 701
DB 1393 CTATGGAGGCTACATCCAGAGGAAGTGGAAAGTCCCTTCTCGGGGTTTCCAGAT 1452
QY 702 PheTyProAsnAsnLeuAsnCysThrTrpIleIleGlnThrSerHisGlyValGlyVal 721
DB 1453 TTTATCCAAACTCTCTAAACGACGCTGACCATTTGAAGTCTCATGGGAAAGAGTT 1512
QY 722 PhePheThrPheHisThrPheHisLeuGluSerGlyHisAspTyThrLeuIleThrGlu 741
DB 1513 CAAATGATCTTTCACACCTTTCATTTGAGAGTTCCACGACTATTTACTGATCAGAG 1572
QY 742 AsnGlySerPheThrGlnProLeuArgGlnLeuThrGlySerArgLeuProAlaProIle 761
DB 1573 GATGGAAGTTTTCGAGACCCGCTGCCAGCTCACCCGGTGGGTTGCTCATACATC 1632
QY 762 SerIleGlyLeuTyArgIleAsnPheThrAlaGlnValArgPheIleSerAspPheSerMet 781
DB 1633 AAGCAGGCTGTGTGGAACTTCACTGCCAGCTTGTTATATGAGACTTCTCAAT 1692
QY 782 SerTyArgGluGlyPheAsnIleThrPheSerGluTyArgLeuGluProCysGluGluPro 801
DB 1693 TCGTACAGAGGCTTCAATATCATTTTCAGAAATATGACTTGGAGCATGTGATGATCT 1752
QY 802 GluValProAlaTySerIleArgLysGlyLeuGlnPheGlyValGlyAspThrLeuThr 821
DB 1753 GGAATCCTCGCTTCAGCCGGAAGATTTGTTTTCATTTGGTGGGAGACTCTGACG 1812
QY 822 PheSerCysPheProGlyTyArgLeuGluGlyThrAlaArgIleThrCysLeuGlyGly 841
DB 1813 TTTTTCGCTTCTCGGAGATATCGTTTGAAGAGTCCGCCAAGCTTACCTCGGGGTGG 1872
QY 842 ArgArgArgLeuTyPheSerSerProLeuProArgCysValAlaGluCysGlyAsnSerVal 861
DB 1873 GCGCGCGGTGTGTGAGTGCACCTCTGCCAAGGTGTGTGCCAATGTGAGAGAACTGTC 1932
QY 862 ThrGlyThrGlnGlyThrLeuLeuSerProAsnPheProValAsnTyArgAsnAsnHis 881
DB 1933 AAGGAAATAGAGAAATTAATTAATCTGTCTCCAAATTTTCCATTCCAATTAATTAATCAAT 1992
QY 882 GluCysIleTySerIleGlnThrGlnProGlyLysGlyIleGlnLeuLysAlaArgAla 901
DB 1993 GAGGTATCTATAAATAGAAACAGAAACCGGCAAGGCACTCCACTTGAACACAGAAAC 2052
QY 902 PheGluLeuSerGluGlyAspValLeuLysValTyArgPGLYAsnAsnAsnSerAlaArg 921
DB 2053 TTCCAGCTGTGTTAAGAGATTAATTAAGGTATATGAGAAAGACAGCTTCTCAGCT 2112
QY 922 LeuLeuGlyValPheSerHisSerGluMetMetGlyValThrLeuAsnSerThrSerSer 941
DB 2113 CCACTGGGCACTTCACTAAATAGAACTTCTGGGGCTGATCTTAAACGACATCCAT 2172
QY 942 SerLeuTyPheAsnPheIleThrAspAlaGluAsnThrSerGlyPheGluLeuHis 961
DB 2173 CACCTGTGCTAGATTCAACCAATGATGTGACACCGACCAAGCTTTCACTCAC 2232
QY 962 PheSerSerPheGluLeuIleLysCysGluAspProGlyThrProLysPheGlyTyArgLys 981
DB 2233 TATACAGTTTGTATCTGTGAAATGTGAGATCCGGGATCTCTAATACTACGGTATAGG 2292
QY 982 ValHisAspGluGlyHisPheAlaGlySerSerValSerPheSerCysAspProGlyTyArg 1001
DB 2293 ATCCGTATGTAAGGCCACTTTACCGACACTGTAGTCTGTACAGCTTGCAACCCGGGTAC 2352
QY 1002 SerLeuAspGlySerGluLeuLeuCysLeuSerGlyGluArgArgThrTrpAspArg 1021
DB 2353 GCCATGATGAGCAGCAACCCCTGACCTGTGTGAGTGGAGACAGGAGATGTGGACAA 2412
QY 1022 ProLeuProThrCysValAlaGluCysGlyGlyThrValArgGlyGluValSerGlyGln 1041
DB 2413 CCACTACCTTCCGTCATAGCGGAATGTGTGTGATTCATTCATGACCAACATCAGACCA 2472
QY 1042 ValLeuSerProGlyTyThrProAlaProGlyTyArgLysAsnLeuAsnCysIleThrThrIle 1061
DB 2473 ATATTTCCCTCGCTATACCTCCTGATGACAAACACCTCCTCAGCTGATATTA 2532
QY 1062 GluValAlaGlyCysThrIleGlyLeuHisPheLeuValPheAspThrGluGluVal 1081
DB 2533 GAGGAGACCCAGAAAGAACATTTAGCTTCCATTTTATTTGACACGGAGATGCT 2592
QY 1082 HisAspValLeuArgIleTrpAspGlyProValGluSerGlyValLeuLeuLysGluLeu 1101
DB 2593 CACGACATCTCAAGGTCTGGAGCGGGCGGTGGACAGATCATCTCTGTAAGAGTGG 2652
QY 1102 SerGlyProAlaLeuProLysAspLeuHisSerThrPheAsnSerValValLeuGlnPhe 1121
DB 2653 AGTGCTCCGCTTCCGAGGACATCCACAGACCTTCACTCACTCACTCACTGAGTTC 2712
QY 1122 SerThrAspPhePheThrSerLysGlnGlyPheAlaIleGlnPheSerValSerThrAla 1141
DB 2713 GACAGGACTTCTTCATAGCAAGTGTGCTTCTCATCCAGTTTCCACCTCAATTGCA 2772
QY 1142 ThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerGlyAspSerTrpGlu 1161
DB 2773 GCACCTGTAAAGATTCAGATATGCCCCAAATGACCCCTATGAGACAGCAGAGAG 2832
QY 1162 AlaGlyAspSerThrValPheGlnCysAspProGlyTyArgAlaLeuGlnGlySerIleGlu 1181
DB 2833 GCTGAGAGACCCGTACATTTCCAGTGTACCTGTGCTATCAGCTTCCAAAGCAACGCCAA 2892
QY 1182 IleSerCysValLysIleGluAsnArgPhePheThrGlnProSerProProThrCysIle 1201
DB 2893 ATCACCTGTGTGACGTGAATTAACCGGTTCTTTGGCAACAGACCTCTCAATGATTA 2952
QY 1202 AlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeuSerProAsnTyPro 1221
DB 2953 GCTGCTTGTGAGGGGAATCTAGACGGCCGACAGGTGTATTTGTGTACCCCACTACCA 3012
QY 1222 GluProTyProProGlyLysGluCysAspTrpLysValThrValSerProAspTrpVal 1241
DB 3013 CAGCGTATCTCTCGGGAAGGAATGTGACTGAGAGTGAAGTGAACCCGGACTTTGTC 3072
QY 1242 IleAlaLeuValPheAsnIlePheAsnLeuGluProGlyTyArgPheLeuHisIleTy 1261
DB 3073 ATCCGCTTGATTAATCAAAAGTTTCACATGAGGCCAGCTATGACTTCTCAACATCTAT 3132
QY 1262 AspGlyArgAspSerLeuSerProLeuIleGlySerPheTyArgLysGlnLeuProGly 1281
DB 3133 GAAGGGAAAGTTTCCAAACAGCCCTCATTTGGAGTTTACAGGGCTTCCAGGCCCA 3192
QY 1282 ArgIleGluSerSerSerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerValSer 1301
DB 3193 AGAATTAAGAGTACCGGAAACAGCTGTCTTCTGGCATTTGAGATGCTCCGTGGGC 3252
QY 1302 AsnAlaGlyPheValIleAspTyThrGluAsnProArgGluSerCysPheAspProGly 1321
DB 3253 CTTTCAGGGTTCGCTATTAATTAAGAGAAACCAAGGGAACCTGTTTGGACCAAGGA 3312
QY 1322 SerIleLysAsnGlyThrArgValGlySerAspLeuLysLeuGlySerSerValThrTy 1341
DB 3313 AATATATGATGAGCAAAAGTGTGAAACGACTTCAAGCTTGGCTTCCACATCACTTAC 3372
QY 1342 TyrCysHisGlyGlyTyArgLysValGluGlyThrSerThrLeuSerCysIleLeuGlyPro 1361
```

```
Db 3373 CAGGTGACCTGCTGCTAATAAGATTCTTGACCCCTCATCATCATCTGTGATTTGGGGCT 3432
Qy 1362 AaegilyyarpovaalTPraenapnProaRgProValCystrhalaproCyeglyglIn 1381
Db 3433 GATGGGAAACCCCTCTGGGACCAAGTGTCTGCTCCCAATGCTCTCTGGAGGGCCAG 3492
Qy 1382 TyrValGlySerAaRgVValValLeuSerProaantYrProGlnaantYrThSerGly 1401
Db 3493 TACCGGATCAGAAAGGGGTTGTTTATCACCACCAATCCCCCAATATACAGAGTGT 3552
Qy 1402 GlnIleCyseuLeuYrPheValThrValProLySaRgTyRValValPheGlyGlnPheA 1421
Db 3553 CAATATGCTCTTATTCATCAGGTACCGGTACCAAGAAATTCGTGCTTTGGACAGTTGGCC 3612
Qy 1422 PhePhehIsthrAlaLeuAenAaRgValValGluValHisAaRgGlyHisSerGlnHisSer 1441
Db 3613 TATTTCCAGACAGCCCTGMAATGATTTGGCAGAAATATTGATGGAACCCATGACAGAGCC 3672
Qy 1442 ArgLeuLeuSerSerLeuSerGlySerPhehIsthrGlyGlySerLeuProLeuValThrSer 1461
Db 3673 AGACTTCTCAGCTACTCTCGGGGCTCAGTCAAGGGAACATTTGCCCTTGCTACGTC 3732
Qy 1462 AaenIValLeuLeuIleYrPheSerAlaYsGlyLeuValProAlaArgGlyPhehIsthr 1481
Db 3733 AATCAAAATTCGCTCCGATTCAGTGCAGAAAGCGGTGCTGCTGCCCGGCTTCACATTC 3792
Qy 1482 ValTyRglnAlaValProaRgThrSerAlaThrGlnCyseuSerSerValProGlnProaRg 1501
Db 3793 GTGATACAGCTGCTCTCGTACCAAGTACACCAATGACAGCTGTGCTCCGACCCAG 3852
Qy 1502 TyRgIlySaRgLeuGlySerAaRgPheSerValGlyAlaIleValArgPheGlyCyAa 1521
Db 3853 TACGAAAGAAATTTGTTCTGAGTTTCTGCCGCTCATGTCGATTCGATTCAGATTCAC 3912
Qy 1522 SerGlyTyRAlaLeuGlnGlySerProGlnIleGlyCyseuProValProGlyAlaLeu 1541
Db 3913 CCGGATACCTCGCTTCAAGGGTTCCAGCGGCTCCAGCTCCAGTCCGAGCCCAAGCCCTG 3972
Qy 1542 AlaGlnTPraenValSerAlaProThrCyValValProCyeglyGlyAaenLeuThGlu 1561
Db 3973 GCACAGTGGAAAGCAGACAGATCCCAAGCTGTGTGATCCCTGACGTGGCAATTTCACTCA 4032
Qy 1562 ArgRgGlyThrIleLeuSerProGlyPheProGlnProTyRLeuAaenSerLeuAaenCy 1581
Db 4033 CGAAAGGATCAATCTGCTCCCGCTACCTGAGCCATGCGAAACCACTTAACCTGT 4092
Qy 1582 ValTyRySileValProGlnGlyAlaGlyIleGlnIleGlnValIaSerPheVal 1601
Db 4093 ATATGGAAGATCATAGTTACGAGGGCTCGGGAATTCAGATCCAAAGTATCATGTTGCC 4152
Qy 1602 ThrGlnIaenTPraSerSerLeuGluValPheAaRgIlyAaAaenThrValThrMet 1621
Db 4153 ACGAGAGCAAACTGGGACTCCCTTGAGATCCAGATGGTGGGATGTGACCGCACCA 4212
Qy 1622 LeuGlySerPheSerGlyThrThrValProAlaLeuAaenSerThrSerAaenGln 1641
Db 4213 CTGGGAAGCTTCTCAGGACCAACAGTACCGGCTGCTGAAACAGTACTCCAAACAATC 4272
Qy 1642 TyrLeuHisPheTyRSerAaRgIleSerValSerAlaIaGlyPhehIsthrGlyTyR 1661
Db 4273 TACCTGCATTTCCAGTCTGACATTAAGTGTGACACTGTGTTTCCACTGGAATACAA 4332
Qy 1662 ThrValGlyLeuSerSerCyPProGlnProAlaValProSerAaenGlyValLeuThrGly 1681
Db 4333 ACTTAGGCTCTGCTGCTGCTCAAGAACCAAGCCCTCCCAACACAGATCAAAATGCGA 4392
Qy 1682 GlnuRgTyRLeuValAaenAaRgValValSerPheGlnCyeglyProGlyTyRAlaLeuGln 1701
Db 4393 GATGGTACATGATGGAAGACAGTCTCTCTTCCAGTGCAGGCCGGGTACACCTGAG 4452
Qy 1702 GlnuHAlaHisIleSerCyMetProGlyThrValArgArgTPraenTyRProPro 1721
Db 4453 GGCGTTCCCAATTTCTGTATATGACGAGGACCGTTGCGGTTGGAATACCTGCTCC 4512
Qy 1722 LeuCyValIleAlaGlnCyeglyGlyThrValGlnGluMetGlnGlyValIleLeuSerPro 1741
Db 4513 CTGTGATTCACACTGTGAGGGACGCTAGACACTTGGGTGTGTGATCTGAGGCC 4572
Qy 1742 GlyPheProGlyAaenTyRProSerAaMetAaRgCySerThrTyRySileAlaLeuProVal 1761
Db 4573 GGCTTCCAGGTTCTTACCCCAACACTTGAAGTACCTGACCTGAGGATCTCATTCACATC 4632
Qy 1762 GlyPheGlyAlaHisIleGlnPheLeuAaenPheSerThrGlnProAaenHisAaRgTyR 1781
Db 4633 GGCTATGTGCATATATTCAAGTTTCTGAATTTTCTACCAAGCTATATCATGACTTCT 4692
Qy 1782 GlnuIleArgAaenGlyProTyRglnThrSerArgMetGlyAaRgPheSerGlySerGln 1801
Db 4693 GAAATTCAAAATGACCTTACACACACAGCCCTGATGTGACAAATTTAGCGGACGAT 4752
Qy 1802 LeuProSerSerLeuLeuSerThrSerHisGlyThrThrValTyRPhenHisSerAaRgHis 1821
Db 4753 CTCCCGCGGCTGCTGAGACACACAGCATGAAACCTCATCCACTTTATAGTACAT 4812
Qy 1822 SerGlnAaRgProGlyPheTyRLeuGlyTyRglnAlaTyRglnLeuGlnGlyCyPro 1841
Db 4813 TCGCAAAACCGGCAAGGATTTAACTTGAACCTTACCAAGCTATGATTAAGAACTGTCA 4872
Qy 1842 AaRgProGlnProPheAlaAaenGlyIleValArgGlyAlaGlyTyRAlaValGlyGlnSer 1861
Db 4873 GATCCACCCCATTTCAAGATGGGTATCATATCAACCGGATTAACCGGGGCAATCA 4932
Qy 1862 ValThrPheGlyCyseuLeuProGlyTyRglnLeuThrGlyHisPProValLeuThrCyGln 1881
Db 4933 GTATCTTTCAGTATTTCTGAGTACATTTCTGATAGCCATCTGTGCTCCTCAGTGTGAG 4992
Qy 1882 HisGlyThrAaRgAaenTPraSerPhePheLeuLeuProLyCyseuValProCyeglyGly 1901
Db 4993 CATGGATCAACAGAACTGGAATCTGATCCCTTTTCAAGATGTGATGCCCTTGGGGTAC 5052
Qy 1902 AaenIleThrSerSerAaenGlyThrValTyRSerProGlyPheProSerProTyRSerSer 1921
Db 5053 AACGTAACTTCTCAAGAACGACACCATCTACTCCCTGCTTCTGTGATGATACCATC 5112
Qy 1922 SerGlnAaRgCyValTPraLeuIleThrValProIleGlyHisGlyValArgLeuAaenLeu 1941
Db 5113 CTGAAGACGTGATTTGGCTCATACCGGTGCTCCAGGGACAGGATTTACATCAATCTC 5172
Qy 1942 SerLeuLeuGlnThrGlnProSerGlyAaRgPhehIsthrIleTPraPrgIyProGlnGln 1961
Db 5173 ACCCTGTTCAGACGGAAGCTGTCAAGATTAATGCTGTGTGGAGCGGTCCGATCAG 5232
Qy 1962 ThrAlaProArgLeuGlyValPheThrArgSerMetAlaLySlyThrValGlnSerSer 1981
Db 5233 AACCTACCCCAAGCTGGAGTGTTCAGTGGCAACACCCCTCGAAGAGGGGTATAGTCTC 5292
Qy 1982 SerAaenIValLeuLeuLySlyPhehIsthrArgAaRgAlaIaThrGlyGlyIlePheAlaIle 2001
Db 5293 ACCAACCAAGTCTGCTCAAGTGTCCACAGGACTTTTCAATGAGGCTTCTTGTCTC 5352
Qy 2002 AlaPheSerAlaTyRProLeuThrTyRCyPProProThrIleLeuProAaenAlaGln 2021
Db 5353 AATTTCCAGCATTTTCAAGTCAAGAAATGTCAACCTCCCAAGCGTTTCCACAGGACGAA 5412
Qy 2022 ValValThrGlnAaenGlnGlyPheAaenIleGlyAaRgIleValArgTyRArgCyseuPro 2041
Db 5413 ATGCTTACTAGATGATGATTTGAGATGAGAAATTTTGAAGTACCAAGTCCAGTCCACCC 5472
Qy 2042 GlyPheThrLeuValGlyAaenGlnIleLeuThrCyValSlySerLeuGlyTyRLeuGlnPhe 2061
Db 5473 GGGTACCTTGTGGGAGACGACATTTCTGACTTTCAGTCACTTCCAGTTGCACTTT 5532
Qy 2062 GlnuGlyProProIleCyeglyValHisCyseProThrAaenGlnLeuThrAaRgSer 2081
Db 5533 GAGGTTCTCTCCCAATGTGAAGACAAATGCCCAATGAAATGACCGGACTGATCA 5592
```

```

QY 2082 ThrGlyValIleLeuSerGlnSerTyrProGlySerTyrProGlnPheGlnThrCysSer 2101
DB 5593 TGGGAGCATATTCAGTCAGCCAGGTAATCGGGTAATATTTAACTCCCAAGCTGGCTCT 5652
QY 2102 TTPLeuValArgValGlnProAspTyrAsnIleSerLeuThrValGlnTyrPheLeuSer 2121
DB 5653 TGGAGATTAATAGTGAACCAAACTAACATTACCATCTTGTGGACACATTTCCAAAGT 5712
QY 2122 GlnLysGlnTyrAspGlnPheGlnIlePheAspGlyProSerGlyGlnSerProLeuLeu 2141
DB 5713 GAAAAGCACTTGAATGACATGGAAGGTGTTGATGGTCTTCCTGGCAAAAGCTCTGCTGA 5772
QY 2142 LysAlaLeuSerGlyAsnTyrSerAlaProLeuIleValThrSerSerAspSerVal 2161
DB 5773 GAACTTAAAGTGGAGATCATCTGACATCAATTAATTTTACAGCAGAGATTAACATGTA 5832
QY 2162 TyrLeuAspGlyPheSerSerAspHisAlaTyrAsnArgLysGlyPheLysIleArgTyrSer 2181
DB 5833 TATCTCCGCTGGTCCATGACATGCACTGCAACAGTAAGATTCAGATTCGCTATGCA 5892
QY 2182 AlaProTyrCysSerLeuProArgAlaProLeuHisGlyPheIleLeuGlyGlnThrSer 2201
DB 5893 GCACTTACTGCAAGTTTACCCACCCCTGAAGAAATGGGGTATTTAAACAGAGACTGCA 5952
QY 2202 ThrGlnProGlyGlySerIleHisPheGlyCysAsnAlaGlyTyrArgLeuValGlnHis 2221
DB 5953 GGAGCGCTTGGAGACAAAGTGCATTAATTTTTCGAAGCCTGGATACCAATGGTGGCCAC 6012
QY 2222 SerMetAlaIleCysThrArgHisProGlnIlyThrIleLeuTyrPheSerGlnAlaIlePro 2241
DB 6013 AGCATATGACCTGTGAACGAAACCCACTTGGCATGTAACAGTGGAGACTCCCTCACCCCA 6072
QY 2242 LeuCysGlnAlaLeuSerCysGlyLeuProGlnAlaProLysAsnGlyMetValPheGly 2261
DB 6073 CTGCGCAGCGCTGGTCTCTGTGAATCCCAATCCCAAGAAACGGTTTACATTACCGG 6132
QY 2262 LysGlnIlyThrValGlnIlyThrLysAlaValTyrSerCysSerGlnGlyTyrHisLeuGln 2281
DB 6133 AACAGTTCACTTGGACAGTAAGTGGTCTATGATGATGATGAGGGCTTCAAGCTTGA 6192
QY 2282 AlaGlyAlaGlnAlaThrAlaGlnCysLeuAspThrGlyLeuTyrPheSerAsnAlaGlnVal 2301
DB 6193 TCCAGCCAGCAGACAAACAGCCGTGTCAAGAAATGGGCTGTGGAGTAACAAAGGGGAG 6252
QY 2302 ProProGlnCysValProValThrCysProAspValSerSerIleSerValGlnHisGly 2321
DB 6253 CCGCCCAAGTGAAGCCGGTCCGCTGGCCCAAGCATTAAGCTCAGACTCAGAACTATGTC 6312
QY 2322 ArgThrPheLeuIlePheGlnIlyThrGlnIlyThrPheGlnAlaGlnLeuMetLeuIleCys 2341
DB 6313 ATCTGGAGGCTGGTTCAGAGTCCCTGTAATGATGATGATGATGATGATGATGATGATGAT 6372
QY 2342 AspProGlyTyrTyrTyrTyrThrGlyGlnArgValIleArgCysGlnAlaAsnGlyLys 2361
DB 6373 AGTCCTGTACTACTTAAGAGAGCTGGAGGCTCTCGCGGTCCAGGCCAATGGAGCTGG 6432
QY 2362 SerLeuGlyLysSerThrProThrCysArgIleIleSerCysGlyGlnLeuProIlePro 2381
DB 6433 AACATAGGAGATAGAGGCCAAGCTGTGAGTTATCTCGTGGAGAACCTTTCCTTCC 6492
QY 2382 ProAsnGlyHisArgIleGlyThrLeuSerValTyrGlyAlaThrAlaIlePheSerCys 2401
DB 6493 CCAAAATGGCAACAAGATTTGAACGTTTAAAGTTAATGGGCAACAGCTAATATTTACGTGC 6552
QY 2402 AsnSerGlyTyrThrLeuValGlySerArgValArgGlyCysMetAlaAsnGlyLeuTyr 2421
DB 6553 AACACGGGCTACACGCTTGGGGTCTCATGTCAGAGATGCTTGGCAAAATGGGCTCTGG 6612
QY 2422 SerGlySerGlyValArgCysLeuAlaGlyHisCysGlyThrProGlnProIleValAsn 2441
DB 6613 AGCGGAGCGAAACTGATGTCTGGCTGGCACTGGCGTTTCCCAAGACCCGATGTGTGAAC 6672

```

```

QY 2442 GlyHisIleAsnGlyLysLeuTyrSerTyrArgGlySerValValTyrGlnCysAsnAla 2461
DB 6673 GGTCACTATTAAGTGAATGAGCTTCAAGTATCAGAGACACAGGTGGTTTACCAGTCAATCT 6732
QY 2462 GlyPheArgLeuIleGlyMetSerValArgIleCysGlnGlnAspHisIleTyrPheSerGly 2481
DB 6733 GGTTCGCGCTTGTGGAACTTCCGTGAAGATATGCTCAAGACACCAAGTGGTCTGGA 6792
QY 2482 LysThrProPheCysValProIleThrCysGlyHisPheProGlyAsnProValAsnGlyLeu 2501
DB 6793 CAACCGCTGTGTGTGCCCATCAATGTGGTCACTCGAAACCTCCAGCCGATTC 6852
QY 2502 ThrGlnGlyAsnGlnPheAsnLeuAsnArgValValLysPheValCysAsnProGlyTyr 2521
DB 6853 ACTAATGGAAGTGAAGTTCAACTGATATGTCGTAATTTTCACTCGCAACACGGGCTAT 6912
QY 2522 MetAlaGlnGlyAlaAlaArgSerGlnCysLeuAlaSerGlyGlnThrSerAspMetLeu 2541
DB 6913 TTGCTCAGGCGGTGTCTGACCCAGCTGTCCGAGCAGACGCGCAGTGGAGTCCCTCTG 6972
QY 2542 ProThrCysArgIleIleAsnCysThrAspProGlyHisGlnGlnAsnSerValArgGln 2561
DB 6973 CCCACGTGTGAGTGTGAATCTGTGATCCAGGCTTGTGGAAATGCCATTCGTAC 7032
QY 2562 ValHisAlaSerGlyProHisArgPheSerPheGlyThrThrValSerTyrArgCysAsn 2581
DB 7033 GGGCAACAGAACTTCCCTGAGAGTTTGAATGGAATGAGATATCTGTACATTGCAAG 7092
QY 2582 HisGlyPheTyrLeuLeuGlyThrProValLeuSerCysGlnGlyAspGlyThrTyrAsp 2601
DB 7093 AAGGATTTTACTTCTGGATCTTCAAGCTTCAAGCTTGAAGTGAAGGCTTATAGGAGC 7152
QY 2602 ArgProAspProGlnCysLeuLeuValSerCysGlyHisPheProGlySerProHisSer 2621
DB 7153 CGATCCCTGCCAAATGTTTGGCTATATCGTGTGACACCCAGGGGCTCCGCCAACGCC 7212
QY 2622 GlnMetSerGlyAspSerTyrThrValGlyValAlaValArgTyrSerCysAlaGlyLys 2641
DB 7213 GTCCTCATGAGAGCTTTTAACTTAATGAGCCGCTGTGCACTACTCTGCAAGGGAGGC 7272
QY 2642 ArgThrLeuValGlyAsnSerThrArgMetCysGlyLeuAspGlyHisIleTyrThrGlySer 2661
DB 7273 GAGACCTCATAGGACAGACAGACAGAGTGTGCCAGAAACAGTCACTGGAGCGGGCA 7332
QY 2662 LeuProHisCysSerGlyThrSerValGlyValCysGlyAspProGlyIleProAlaHis 2681
DB 7333 CTGCCCACTGACAGGAATTAATCTGTGATTTCTGTGTATTCGGGAGACCCCAAGCAT 7392
QY 2682 GlyIleArgLeuGlyAspSerPheAspProGlyThrValMetArgPheSerCysGlnAla 2701
DB 7393 GGGTCTGGGCTTGGTGAATTAATTAAAGAAAGTCTTCCCGCTTCTCTGTGAATG 7452
QY 2702 GlyHisValIleLeuArgGlySerSerGlyLysArgThrCysGlnAlaAsnGlySerTyrPheSerGly 2721
DB 7453 GGGCAGCAGCTGAAGGGCTCCCTCGAACCGACGTTTGTCTCAATGGGTCATGGTCAAGA 7512
QY 2722 SerGlnProGlyCysGlyValIleSerCysGlyAsnProGlyThrProSerAsnAlaArg 2741
DB 7513 CTGACAGCGGTGTGAGGCGGTCTCTGTGGCAACCTGTGCACACCCACACAGCGAATG 7572
QY 2742 ValValPheSerAspGlyLeuValPheSerSerSerIleValTyrGlyCysArgGlnGly 2761
DB 7573 ATTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7632
QY 2762 TyrTyrAlaThrGlyLeuLeuSerArgHisCysSerValAsnGlyThrThrThrGlySer 2781
DB 7633 TACAAGACTCGAGGCTCATGACACGCGCATTTGACACGCAATGGAGCTGGACAGGACT 7692
QY 2782 AspProGlyLysLeuValIleAsnCysGlyAspProGlyIleProIleAsnGlyLeuArg 2801
DB 7693 GCTCCGACACTGACAAATTAATGTTGGGAGATCCAGGCACTATGCAATATGCAATCCAG 7752
QY 2802 LeuGlyAsnAspPheArgTyrAsnLysThrValThrTyrGlnCysValProGlyTyrMet 2821

```

Db	Accession	Gene	Protein	Length	Source
Db	7753	TTGGGACCGACCTTACCTTCAACAGACTGTGACCTTACGTGTAACTCCAGGCTATGTC		7812	
Qy	2822	MetGlnSerH1sArgValSerValLeuSerCysThrIlyAspArgThrTrpAsnGlyThr	2841		
Db	7813	ATGGAAAGACATCATCCGCCACTATTCGGCTGTACCAAGACGGAGGTGAAATCCAGAC	7872		
Qy	2842	LysProValCysGlyAlaLeuMetCysAlaBProProLeuIleProAsnGlyLysVal	2861		
Db	7873	AAACCTGCTCTCCAAAGCCGGTGTGTGTCTCAGCCGCCGCCGCGCAAAATGGAAACAGTG	7932		
Qy	2862	ValGlySerAspPheMetTrpGlySerSerValThrTrpAlaCysLeuGlnGlyTrpGln	2881		
Db	7933	GAGGGAATGATTTTCCTCGCTGGGGCTCCAGCAATGATTAACACTGATGAGACGGTTAACAG	7992		
Qy	2882	LeuSerLeuProAlaValPheThrCysGlnGlyAsnGlySerTrpThrGlyLysLeuPro	2901		
Db	7993	CTCTCTCACTCGCGCATCTCTCTCGTGAAGTCCGGGGGTGTGGAAAGAGATATCCCC	8052		
Qy	2902	GlnCysPheProValPheCysGlyLysAspProGlyValProSerArgGlyLysArgGlnAsp	2921		
Db	8053	CAGTGTCTCCCTGTGTGTTCGGCGAGACCTGTGACATCCCGGAGAAAGCGCATTAATGCGG	8112		
Qy	2922	ArgGlyPheSerTrpArgSerSerValSerPheSerCysHisProProLeuValLeuVal	2941		
Db	8113	AAAGATTTCACCTTAATAGTCCGAAAGTCTTTCAGTCAATCCAAATCTCAATTAATCTGTG	8172		
Qy	2942	GlySerProAlaArgPheCysGlnSerAspGlyThrTrpSerGlyThrGlnProSerCys	2961		
Db	8173	GGATCTCTCCAAAGAGTCTGCCAGCTACCGGACGTGGAGCGCATTAACCACTACCTGC	8232		
Qy	2962	IleAspProThrLeuThrThrCysAlaAspProGlyValProGlnPheGlyIleGlnAsn	2981		
Db	8233	ATTGATCTCTGATTAACACCTGCGCCAGACCTGTGATGCGCACATTTGGAATACAGAT	8292		
Qy	2982	AsnSerGlnGlyTrpGlnValGlySerThrValLeuPheArgCysGlnIlyGlyTrpLeu	3001		
Db	8293	AGCTTCAGAGCTTAAGAGTGTGAAGACCGGTTTTTTCAGGTGCAGAAAGGCTACCAT	8352		
Qy	3002	LeuGlnGlySerThrThrArgThrCysLeuProAsnLeuThrTrpSerGlyThrProPro	3021		
Db	8353	ATTCAAGGTTCCAGACATCCGACCTGCTTGCCAATTAACATGAGTGGGATACAGAC	8412		
Qy	3022	AspCysValProHisAsnAspArgGlnProGlnThrProThrHisAlaAsnValGlyAla	3041		
Db	8413	GAATGTATACCTCACTGCTGCAGACAGCAGCAAAACCCGGGACACGCGGATGTGAGACC	8472		
Qy	3042	LeuAspLeuProSerMetGlyTrpThrLeuIle-ThrProAlaArgAlaSerProse	3061		
Db	8473	ATCGATCTCTCACTTTCGGCTACCTTAAGTGTACACTGCATCCAGGCTTTTCTCTC	8532		
Qy	3061	ArgValAlaProSerThrAlaProAlaArgArgMetAlaIleGlyGlnAlaSerArg	3080		
Db	8533	GCAGGGGATCTGACACAGAACATGTAAAGCAGCATGAATGACAGAAAGTGGC	8590		
RESULT 7	AX374891	10673 bp	DNA	linear	PAT 01-WAR-2002
LOCUS	AX374891				
DEFINITION	Sequence 1 from Patent WO0210199.				
ACCESSION	AX374891				
VERSION	AX374891.1				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE	1				
AUTHORS	Welcher, A.A. and Elliott, G.S.				
TITLE	C3b/c4b complement receptor-like molecules and uses thereof				
JOURNAL	Patent: WO 0210199-A 1 07-FEB-2002;				
FEATURES	Amgen Inc. (US)				
source	Location/Qualifiers				
	1..10673				

```

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
334..9543
CDS ..

```

protein_id="CA026763.1" /db_xref="GI:19169787" /translation="MSNWMLHLQSDSDIGSPGFKAAYOEIERKGSCGDPGIPAYGKRTGSGSLHGDTLTFECPPAFELVGRVAVITQQOONMGNKBECSVSCSPFASGIIILSPNYEEYGGNMNMCWMLIISESPGRHILIFNDPVEOPDPLAKDGIISDITVLGTFSGNEVPSQLASGHLVRLFEPOSDHSTTGROGNITVTTFGQNEHCDPGIINAGRFGLLSSVSHFCHDDGDFVQKQGESLITCLIDONNVMSVTPRCBAPGDLHSGSVLLPCMGXYKXDSLHCWMLIEAKRGHSKITRPRQTVNYVTLERPDPAASSPLIGEYHGTQAPVLISTGNFMVLLFTTDSRSISGLVILIBSVTLLESBCLDPGIPVNXHRRGDFGQIRSVLSTGCDGYTLSDDEPLVCERNQNMHALSPCALCGGVIQKSGSVLSPGFPDPEPMSLXKWTWITIEVSHKGVDQMFHTFHLLSHDYLITDEDSRESEVALATGSLPHPIKAGLFGNFPAQLRFISDPSIYVEGNIITFSEYDLEDPDPPGAPVARSIGLHFGVSGSLTFPSCFLGRLEGAAXKILCTLCGRGRVMAPLPRCAEAGSAGSNGEGLILSPNPGSDNNHNCICTKIETHEKAGKHLIRSRQLFEGDGLIKGDLKSSSRLGTFTNEILLGIILNSTNHLMLPNTNGSDPDGPOULTYSPDLKCVGDSIPMYGRIIDEGFYTDPVLYLSCNPGVAMGSNTLTCLSDGRVWDXRPLRSCIAECGGQIHAATSGILSGLTPYAADNNHLCWTIEADPGKTIISLHETVDTEMAHILKVMQPVSDILKLEWSSALPEDHISTFNSTLSLQSDSPFISKSQFSIQFSTSLAICNDGMPGNRYSDRSAGDTVPOCDPQYLOQAKITCVQLNRFPMWQDPPICIAACGNLTGPAVILSPYPOPEPKECDMDRKYVNPFDVILIFSPFMERSEYDFILYEEDSNPLIGSGVQAPABEISSGNSLFLAFPSDVSGLSGALIEFKERPRACPDNPINNGTGVDFPKLGSITVQODSGVKTLIDPSITVCVIGADPKPMDVOLSCNAPCGGQVTSAGCVLSPYPHNTAQICLITLTTPKEVYVFGQVPTOPLANDLAEFDGTHAQRLLSLSGHSGETIPLTASQILIRFSAKSGASARGHFVYQAVPRISDQCSVEPRGRIIGSFPSAGSIVAREXNPGPGLQSGSTQSPVNAQAMDNIIPSCVAPGSGNFPORGKTLISPGVEPYGNMNLICWKILITEGSGIQIOWISFTEOMNSLIEHDGVDAPYALGSSGTVPYALINSTNOLYLHFOSDISVAAGHLEKYTVGLAACPEALPMSIKIGRYMNDVLSFOCEPGYTLQGRSHISCMPTVARMNYPSPICATCGTILSTIGVILSFGPQSGPNNLIDCTWRISLPIYGNAHIOFLNSTEANDHLEIDONGFHTSPMISQFSTDIPRALISTTHETLIRHPSDHSQNRQKAYOAYELONPDPNPGNMINSQVYQGSVSECTPGYILIGHVPLTCQGINRMNITPPRCDACGNVITSQNTTISPGPDEYPLKDCIMLITVPEPGHYINFTLLQTEVANDI1AWMDGDNPSOLGVSESGNLALETAVSSTQVLLKFNHSDPNSNGPFAVNLNFAPOLKQCPAPQVQAMLETDDEPFGDPAVYQCHPGYTLVGDILITCKLSQLOFEGESILPTCEBQAPAEVYTSGGSVILSPYPGNVPQCTGMSIKVEPNPNITFVDTFOSEKOPDLAEVDCSSQSPPLVVLNGHTEOSTNFSRNSNOLYRSTHATSKKGFIRVAPAKSICLHPLKNGHILNRTAGVNSKHYHECKPGRMYNGHSMATRCMDPLCMQDLSITPLCOANSCGIPSPSGNCTGNFETLSKVYVECHBEGFKLESSQOATVACOEDISBKNKGPCTCKVACPSIAQISBHTIWRVLSGLINEYGAOVLLSCSPGYILBEMRLRQOANGTNIGDBRSCVATSCGSLSPYPMGNKTIGLTVYGATILPTCNTGYTLVSGHVECLANGTMSGERCLAGHSGDPP1VNGHISGDEFSYRDIIVYQCNPGFRLVGSVILCIODHMSQOTPEVCYITCGHPNPAQFHTSGSEFNLDVNYPTCNTGYLQGVSRACOSNGMSQSPPLPCRVANCSDDPEFVENA1IRHGOONPPSPSEYMSILLYHCKKGYLIGSSALITMANGMLDNLPRCLTASCGHRCVPANAVITGSELPTVGAVYHYSRGESLIGCNTRYCOEDSHMSGALPHCTGNPGECDGPTPAHGSLEDDFTKSLRLRSCMHQOLKSGPRTCLANGSMGLOPYCEAVSVCNPGPTNGMIVSSDGLIFSSSVIYAQWCEKYSGMLTRICLTKQGRMNPASV1ISGCDPGLANGIIOFGDTLTFNKTVSYAACBPYWEAVATIRCTYKGRMNPASVCKVLCPOPVPVNGTEVSGDFPMWGSISYSCMDQVLLSSALISECBRWMGELIYCLPFCGDPGIP1PAEGRSLSGFTYKSEFVQCSPLVLSASRVCADQWMSGIPOTCIPDAHNTCPDGPGRHIGNNSRGVEYVTPRCRKHIGSGITRCLTANLWMS1IORECTPHACORETPAHADVRA1IDLPTREYTLVTPRCBPLKSGSEHRTCAKDMKTGKSPVCKSGKVRANETTKTPYPSDVPFVMSLMKYTYLEKROPATLTYDVPNATSSKRNATSEASVELKTLGITKKEEHLHLKAFQIKQADIVSVSFRENMDGTVVSSGIERGTFPOGDHKGDKGRLEKGTGEOPLANDQSSSHVTSGSSVAAAL1LVPEALITISGAFYLYLKHRTPRKVQYVNGAGHENSQASFENMYPDNLKTEAKVRDPTTLNVCIVV"

ORIGIN	
Alignment Scores:	
Pred. No.:	0
Score:	10737.00
Percent Similarity:	78.37%
Best Local Similarity:	63.60%
Query Match:	63.21%
DB:	6
US-10-016-248-2 (1-3104) x AX374891 (1-10673)	
	length: 10673
	Matches: 1864
	Conservative: 433
	Mismatches: 567
	Indels: 67
	Gaps: 7

QY	511	Cys61YxL1YH4leuThrIleuSerProSerGlyYthIleuLeuSerProGlyYTPProGlyPhe	530
Db	1120	TGtGTGGACATCTGACAGCGCTCCAGGGAGTCATTTTGCTCTCTGAAATGGCAGAGATAT	117
QY	531	TyrlYasAPAlaleuSerCysAlaTPValIleGluAlaGlnProGlyYTPProIleYs	550
Db	1180	TATAGAGATCTTTACTTGTAATGGATATATGAAGCAAAACAGGCCACTGTATCAAA	123
QY	551	IleThrPheAPArghPheYstYrGluValAsnTYrAspThrLeuGluValArgAspGly	570
Db	1240	ATAACTTTGACAGATTTGACAGAGGGCAATTAAGACACCTTGAGGTGCAGATGGG	129
QY	571	ArgThrTySerAlaProLeuIleGlyValTYrHISGlyYthGlnValProGlnPheLeu	590
Db	1300	CCAGCCAGTTTCGTCCCACTGATCCGCGAGTACAGAGCAACCAGGCCACCCAGTTCTTC	135
QY	591	IleSerThrSerAsnTYrLeuTyrlleuLeuPheSerThrAspLYsSerHISerAspIle	610
Db	1360	ATAGAGACCGGGAACTTCATGTACATCCGTCAATTCAACAATGACAAAGCGCTCCAGCATC	141
QY	611	GlyPheGlnleuArgYrGluThrHISerHISerAspHISerCysleuAspProGly	630
Db	1420	GGTTTCTCATCAGCATATAGAGTGGAGGCTTGAGTGGATTCCTGCTGGACCCGGGC	147
QY	631	IleProValAsnGlyGlnArgHISGlyAsnAPheTYrValGlyAlaValuAlThrPhe	650
Db	1480	ATCCCTGGAACGRCCATCGCCAGCGTGGAGCTTTGGCATCAGGTCCCACTGACTTTC	153
QY	651	SerCysAPAspSerGlyTYrThrleuSerAPGlyGluProLeuGluCysGluProAPhe	670
Db	1540	AGCTGTGACCCGGGGTACACATTAAGTACAGAGCCCTCGTGTGAGAGGAACAC	159
QY	671	GlnTPSerArgAlaleuProSerCysGluAlaleuCysGlyGlyPheIleGlnGlySer	690
Db	1600	CAGTGAACACAGCGCTTGCCAGCTGCAGACGCTCATATGGAGGCTTAATCCAAAGCAAG	165
QY	691	SerGlyThrIleLeuSerProGlyPheProAPheTYrProAsnAsnleuAsnCyArThr	710
Db	1660	AGTGAACAGCTCTTCTCTCGGGTTTCAAGATTTTTCMAACTCTCTAAACGTCAGCAG	171
QY	711	TrpIleIleGluThrSerHISGlyYsGlyValPhePheThrPheHISerThrPheHISer	730
Db	1720	TGGACCATTTGAAGTCTCTCATGGAGAAAGATTCAAAATGATCTTCAACCTTTCATCTT	177
QY	731	GluSerGlyYHISAPTYrleuLeuIleThrGlnAsnGlySerPheThrGlnProLeuArg	750
Db	1780	GAGAGTTTCCCAAGCTATTTCATATCAAGAGATGAAGTATTTTCCGAGCCCGTTGCC	183
QY	751	GlnleuThrArgYserArgLeuProAlaProIleSerAlaGlyLeuTYrGlyAsnAPheThr	770
Db	1840	AGGCTCACCGGAGTGGTGTCTCTCATACGATCAAGAGAGGCTGTTTGGAAACTTCACT	189
QY	771	AlaGlnValArgPheIleSerAPSPheSerMetSerTYrGlnGlyPheAsnIleThrPhe	790
Db	1900	GCCGACGTTTCGTTTATATACGACTTCTCAATTTCGATCAAGGGCTTCAATATCACTTT	195
QY	791	SerGluTYrAspLeuGluProCysGluGluProGluValProAlaTYrSerIleArgLYs	810
Db	1960	TCAGAAATATGACCGAGGACATGATGAATCTTGAGATCCCTCGCTTCACGCCGAAGATTT	201
QY	811	GlyleuGlnPheGlyValGlyAspThrLeuThrPheSerCysPheProGlyTYrArgLeu	830
Db	2020	GGTTTTCACCTTGGTGTGGAGACTCTCTGACGTTTTTCCGTCTCTCGGATATACGTTTA	207
QY	831	GluGlyThrAlaArgIleThrCysleuGlyGlyAlaArgIleArgleuThrPserSerProLeu	850
Db	2080	GAAAGTGCCRCMAAGCTTACCTGCTGGTGGGGGCGCGCTGTGTGGAATGACACCTCTG	213
QY	851	ProArgCysValAlaGlnCysGlyAsnSerValThrGlyThrGlnGlyThrleuLeuSer	870
Db	2140	CAAAGGTGTGTGGCCGATGTGAGCAAGTGTAAAGAAATGAAGAAATTAATTACTGTCT	219
QY	871	ProAsnIleProValAsnTYrAsnAsnAsnHISGlyIleTYrSerIleGlnIleThrGln	890


```
Db 2200 CCAATTTTCATCCAAATATGATTAATACCATGAGTGTATCTTAATAAATGAAACAGAA 2259
Qy 891 ProglYlybglYlleglnleuYalAargalAphelguleuSerGlYlyAspValleu 910
Db 2260 GCCCGCAAGGCGATCCACCTTAGAAACAGAGCTTCAGCTGTTTGAAGGAAATCTCTA 2319
Qy 911 LysValTYrAspGlyAAsnAsnSerAlaArgleuSglYValPheSerHisSerGlu 930
Db 2320 AAGGATATATGATGAAAAAGACAGATTCTCAGCTCCACCTGAGCGACGTTCAAAAAATGAA 2379
Qy 931 MetMetGlyValThrLeuAsnSerThrsSerSerleuThrPleuAspPheIleThrAsp 950
Db 2380 CTCTGGGGCTGATCCTTAACAGACATCAATACCTGTGCTAGAGTTCAACACCAAT 2439
Qy 951 AlaIleuAnthrSerLySglYpHeGluleuHisPheSerSerPheGluleuIleLyS 970
Db 2440 GGATCTGACCGACCGCAAGGTTTTCACCTCACTATACAGTTTGTGATCTGTAATAATGT 2499
Qy 971 GluAspProglYThrProLyPheGlyTYrLySValHisAspGluGlyHisPheAlaGly 990
Db 2500 GAGGATCCGGGACATCCCTTAATCTAGCGCTATAGATCCGTATGAAAGCCACTTAAACGAC 2559
Qy 991 SerSerValSerPheSerCyAspAspProglYTYrSerleuAarglySerGluGluLeu 1010
Db 2560 ACTGATAGTCTGTACAGTTGACACCGGGGTACCGCATGGCAGACACACCTGACC 2619
Qy 1011 CysLeuSerGlyGluAarglyThrTrpAspArgProLeuProThrCysValAlaGluCys 1030
Db 2620 TGTTTGAGTGAGACAGAGAGTGTGGGACAAACCACTGCTGTCATGCGGAAATGT 2679
Qy 1031 GlYglYThrValAarglyGluValSerGlyValleuSerProglYTYrProAlaPro 1050
Db 2680 GGATGTCAGATTCATGACAGCAACATCAGACGAATATGTCCCTGGCTATCCAGCTCG 2739
Qy 1051 TYrGluHisAenLeuAenCyAlleTrpThrIleGluAlaGluAlaGlyCysThrIleGly 1070
Db 2740 TATACAAACAACCTCCACTGCACTGCACTGATTAAGAGCAGACCCAGAAACCACTTAGC 2799
Qy 1071 LeuHisPheLeuValPheAspThrGluGluValHisAspValLeuAarglyLeuAspGly 1090
Db 2800 CTCCATTTTCATGTTTTCGACACGAGATGCTCAGACATCTCCAGAGTCTGGGACCGG 2859
Qy 1091 ProValGluSerGlyValleuLeuLySgluleuSerGlyProAlaLeuProLyAspLeu 1110
Db 2860 CCGGTGACAGTACATCTCTGTAAGAGTGAAGTGGCTCCGCCCTTCGGAGACATC 2919
Qy 1111 HisSerThrPheAsnSerValValleuGlnPheSerThrAspPheThrSerLySglN 1130
Db 2920 CACAGCACTTCACATCACTCACTCGAGTTTCAGACGACGACTTCTTATACGACAGTCT 2979
Qy 1131 GlyPheAlaIleGlnPheSerValSerThrAlaThrSerCysAsnAspProglYIlePro 1150
Db 2980 GGGCTTCCATCCAGTTCTCCACTCAATTCAGCCACCTGTAACGATCCAGATATGCC 3039
Qy 1151 GluAsnGlySerAargSerGlyAspSerTrpGluAlaGlyAspSerThrValPheGlnCys 1170
Db 3040 CAAATAGCACCCCTATGAGACAGACAGAGGTGAGACACCGTCACATTCACAGT 3099
Qy 1171 AspProglYTYrAlaLeuGlnGlySerAlaIleSerCysValIleGluAsnArg 1190
Db 3100 GACCTGGCTATCAGCTCCAGAGACAGCAAAATCACTGTGTGACACTAAATACCGG 3159
Qy 1191 PhePheTrpGlnProSerProProThrCysIleAlaProCysGlyGlyAspLeuThrGly 1210
Db 3160 TTCTTTTGGCAACAGACCTCTCAATAGCATAGCTGTGTGAGGGAATCTGACGGCG 3219
Qy 1211 ProSerGlyValIleLeuSerProAsnThrProGluProGlyrrProProGlyLyGluCys 1230
Db 3220 CCACAGGATGTTATTTGTCAACCAATACACACGCGTATCTCCCTGGGAAAGAAATGT 3279
Qy 1231 AspTrpLySValThrValSerProAspTYrValIleAlaLeuValPheAsnIlePheAsn 1250
Db 3280 GACTGAGAGTAAAAAGTGAACCCGCACTTGTCTACCTGCTGATATTCAAAAGTTTCAAC 3339
Qy 1251 LeuGluProglYTYrAspPheLeuHisIleTYrAspGlyAargAspSerLeuSerProLeu 1270
Db 3340 ATGAGGCCAGCATATGACTTCTTACATATATAGAGGGAAGATTCACACGCCCTTC 3399
Qy 1271 IleGlySerPheTYrGlySerGlnLeuProglYArgIleGluSerSerAsnSerleu 1290
Db 3400 ATTGGAGTTACAGAGGCTCTCAGCCCCCAGAAAGAAATAGAGATACCGGAAACGCTG 3459
Qy 1291 PheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPheValIleAspTYrThr 1310
Db 3460 TTCTGGCATTTCCGATGATGATCCCTCGTGGGCTTCAAGGTTCCGCAATGAATTTAAA 3519
Qy 1311 GluAsnProArgGluSerCysPheAspProglYSerIleuYAsnGlyThrArgValGly 1330
Db 3520 GAGAAACACAGGAGGAGCTTTGTTGACCCAGAAATATATGAATGGAGACAAAGTTGGA 3579
Qy 1331 SerAspLeuYsValleuGlySerSerValThrTYrTYrCysHisGlyGlyTYrGluValGlu 1350
Db 3580 ACAGACTTCAGAGCTTGCTCCACATCACTACAGATGTACCTGGCTATTAATCTT 3639
Qy 1351 GlyThrSerThrLeuSerCysIleLeuGlyProAspGlyLySProValTrpAsnAspPro 1370
Db 3640 GACCCCTCATCATCAGCTGTGTATGGGCTGATGGGAAACCTCCTGGAGCAAGTG 3699
Qy 1371 ArgProValCysThrAlaProCysGlyGlyGlnTYrValGlySerAspGlyValValleu 1390
Db 3700 CTGGCCCTTCGCAATGCTCCCTGTGGAGCCAGATACAGGATAGAAAGGGGTGTTTA 3759
Qy 1391 SerProAsnTYrProGluAsnTYrThrSerGlyGlnIleCysLeuTYrPheValThrVal 1410
Db 3760 TCACCAACTACCCCTCAATTAACACAGCTGTGTAATAGCTCTTATTCATACAGGTA 3819
Qy 1411 ProLyAspTYrValAlaPheGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspVal 1430
Db 3820 CCAAGGAATTCGTTGCTTGGACAGTTGGCTATTTCCAGACGCCCTGAATGATTTG 3879
Qy 1431 ValGluValHisAspGlyHisSerGlnHisSerArgLeuLeuSerSerLeuSerGlySer 1450
Db 3880 GCAGAAATTAATTTGATGAAACCCATGCAAGGCCAGACTTTCACCTCACTCCGGGCT 3939
Qy 1451 HisThrGlyGluSerLeuProLeuAlaThrSerAsnGlnValleuIleYsPheSerAla 1470
Db 3940 CACTCAGGGAAATATGCTTGGCTAGCTCAATCAATTAATTCCTCCGATTAAGTGCA 3999
Qy 1471 LysGlyLeuAlaProAlaAarglyPheHisPheValTYrGlnAlaValAlaProArgThrSer 1490
Db 4000 AAGAGGGGTGCTGTGCCGGGCTTCACCTGTGTATCAAGTGTTCCTCGTAACAGT 4059
Qy 1491 AlaThrGlnCysSerSerValProGluProArgTYrGlyAargLeuGlySerAspPhe 1510
Db 4060 GACACCCCAATGAGCTGTGCTCCGAGCCAGATACGGAAGGAATGTGTTCTGAGAGTT 4119
Qy 1511 SerValGlyAlaIleValAargPheGluCysAsnSerGlyTYrAlaLeuGlnGlySerPro 1530
Db 4120 TTCCGCGCTTCATCGTCCGATTCGATTCGATTCACCCGGAATTCCTGCTTCAAGGTTTCA 4179
Qy 1531 GluIleGluCysLeuProValProglYAlaLeuAlaGlnTrpAsnValSerAlaProThr 1550
Db 4180 GCGCTCAACGCGCAGTCCGAGCCCAACAGCGCTTGCAAGTGAACGACAGATCCAGC 4239
Qy 1551 CysValValProCysGlyGlyAsnLeuThrGluAargArglyTYrIleLeuSerProglY 1570
Db 4240 TGTGTGTATCCCTGCAATGCAATTTCACTCAACGGAAGAGTCAATCTGTCTCCCGCGC 4299
Qy 1571 PheProGluProTYrLeuAsnSerLeuAsnSerValTrpLySValIleValAlaProGluGly 1590
Db 4300 TACCTTAAGCATACGGAACAACTTAAGTGTATGAGAAATGATCATAGTTACGAGAGGC 4359
Qy 1591 AlaGlyIleGlnIleGlnValAlaSerPheValThrGluGlnAsnTrpAspSerLeuGlu 1610
Db 4360 TCGGGAATTCAGATCCAAAGTATGATGTTTGGCCAGGAGCAAGATCTGGACTTCCGAG 4419
```

```

QY 1611 ValPheAspGlyAlaAspAsnThrValThrMetLeuGlySerPheSerGlyThrThrVal 1630
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4420 ATCCAGATGCTGGGGATGTGACCGACCCAGACCTGGAGAGCTTCTCAGGACCCACAGTA 4479
QY 1631 ProAlaLeuLeuAsnSerThrSerAsnGlnLeuTyrLeuHisPheTyrSerAspIleSer 1650
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4480 CCGGACACTGCTGAACGATCTTCCAACCAACTCTACATTCCTGCACTGACATTTAGT 4539
QY 1651 ValSerAlaAlaGlyPheHisLeuGlnTyrLysThrValGlyLeuSerSerCysProGlu 1670
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4540 GTGGAGCTGCTGCTTCCACCTGGAAATACAAACTGTAGGCTTGTGCTGATGCCAAGAA 4599
QY 1671 ProAlaValProSerAsnGlyValLysThrGlnGlyLysValAsnAspValVal 1690
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4600 CCAGCTCTCCACGACACAGATCAAAATCGGAGATCGATACATGTGAACAGAGCTGCTC 4659
QY 1691 SerPheGlnCysGlnProGlyTyrAlaLeuGlnGlyHisAlaHisIleSerCysMetPro 1710
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4660 TCCTTCCAGTGGAGCCCGGGTACACCTCGAGGGCCGTTCCACATTTCTCTGATGCCA 4719
QY 1711 GlyThrValArgAspGlyTrpAsnTyrProProProLeuCysIleAlaGlnCysGlyGlyThr 1730
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4720 GGGACCTTCCGCTGGAGACTATCCGCTCCCTGTGCATTTGCMAACCTGTGGAGGGACG 4779
QY 1731 ValGlnGlnMetGlnGlyValIleLeuSerProGlyPheProGlyAsnTyrProSerAsn 1750
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4780 CTGAGACACTTGGGGTGGTGTATCTGTACGCCCGGCTTCCAGGTTCTTACCCCAACAC 4839
QY 1751 MetAspCysSerTrpLysIleAlaLeuProValGlyPheGlyAlaHisIleGlnPheLeu 1770
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4840 TTAACTGACACTGAGAGATCTCATTCACCATCGGCTATGTGTGACATATTCAAGTTCTG 4899
QY 1771 AsnPheSerThrGlnProAsnHisAspTyrIleGlnIleArgAsnGlyProTyrGlnThr 1790
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4900 AATTTTCTACCGAAGCTATCATGACTCTTCTTAAATTCAAATGAGACCTTACCAACC 4959
QY 1791 SerArgMetGlnGlyArgPheSerGlySerGlnLeuProSerSerLeuLeuSerThrSer 1810
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4960 AGCCCAATGATGAGCAATTGACGGGACAGATCTCCCGCGGCGCTGCTGACACACAG 5019
QY 1811 HisGlnThrThrValTyrPheHisSerAspHisSerGlnAsnArgProGlyPheLysLeu 1830
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5020 CATGAACCCCTCATCTCCCTTTATATGACCATTTGGAAAAACCGGAAAGGATTTAAACT 5079
QY 1831 GlnTyrGlnAlaTyrGlnLeuGlnGlnCysProAspProGlnProPheAlaAsnGlyIle 1850
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5080 GCTTACCAAGCTATGATTAACAGAACTGTCCAGATCCACCCCATTTGAGATGGGTAC 5139
QY 1851 ValArgGlyValaGlyTyrAsnValGlyGlnSerValThrPheGlnCysLeuProGlyTyr 1870
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5140 ATATGCAACTCGGATTAACAGCGTGGGCAATCAGTATCTTTCCAGTGTATCTCGGTAC 5199
QY 1871 GlnLeuThrGlyHisProValLeuThrCysGlnHisGlyThrAsnArgMetTrpAspHis 1890
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5200 ATTCTAATAGCCATCTGTCTCTCATCTTGTACAGATGGGATCAACGAAACTGGAATC 5259
QY 1891 ProLeuProLysCysGlnValProCysGlyGlyAsnIleThrSerSerAsnGlyThrVal 1910
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5260 CCTTTTCCAGAGTATGATGCCCTTGTGGGTACAACTGAATCTTCCAGAACGGACACAC 5319
QY 1911 TyrSerProGlyPheProSerProTyrSerSerSerGlnAspCysValTrpLeuIleThr 1930
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5320 TACTCCCTGGCTTCTCTGATGAGTATCCGATCTCGAAGAGACTGCAATTTGGCTCATACG 5379
QY 1931 ValProIleGlyHisGlyValaArgLeuAsnLeuSerLeuLeuGlnThrGlnProSerGly 1950
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5380 GTGCTCCAGGGAGGAGATTACATCACTTCACTTTCACAGCGAAGACCTGTCAAC 5439
QY 1951 AspPheIleThrIleTrpAspGlyProGlnGlnThrAlaProArgLeuGlyValPheThr 1970
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5440 GATTACATTGCTGTTGGAGCGTCCGATCAGAACTCACCCAGCTGGAGATTTTCACT 5499

```

```

QY 1971 ArgSerMetAlaLysThrValGlnSerSerAsnGlnValLeuLeuLysPheHis 1990
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5500 GGCAACACAGCCCTCCAAACGGCGTATAGTCCACCAACCAAGTCTGCTCAAGTTCCAC 5559
QY 1991 ArgAspAlaAlaThrGlyGlyIlePheAlaIleAlaPheSerAlaTyrProLeuThrLys 2010
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5560 AGGACCTTTCAAATGAGAGCTTCTTGTCCCAATTTCCAGCATTTGAGCTCAAGAA 5619
QY 2011 CysProProThrThrIleLeuProAsnAlaGlnValValThrGlnAsnGlnGlnPheAsn 2030
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5620 TGTCAACTCCCCACAGGTTCCACAGGCAAGAAATCTTATGAGAGATGATGATTCGAG 5679
QY 2031 IleGlyAspIleValArgTyrArgCysLeuProGlyPheThrLeuValGlyAsnGlnIle 2050
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5680 ATAGGAGATTTGTGAAATGACAGATGCCACCCCGGGTACACTTGTGGGAGCCGACAT 5739
QY 2051 LeuThrCysLeuLeuGlnTyrThrTyrLeuGlnPheGlnGlyProProProIleCysGlnVal 2070
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5740 CTGACTTGCAAGCTCATTTCCAGTTGACGTTGAGGGTCTTCTCCCAACATGTAMGCA 5799
QY 2071 HisCysProThrAsnGlnLeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTyr 2090
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5800 CATGCCCGCAAAATGAGAGTCCGAGCTGGATCATGGAGATCATTTCTACGTCCAGGAT 5859
QY 2091 ProGlySerTyrProGlnPheGlnThrCysSerTrpLeuValArgValGlnProAspTyr 2110
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5860 CCGGGATTAATTTTAATCTCCAGACTGCTCTTGAGATATTAAAGTGGAACCAAACTAC 5919
QY 2111 AsnIleSerLeuThrValGlnTyrPheLeuSerGlnLysGlnTyrAspGlnPheGlnIle 2130
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5920 AACATTACATCTTGTGTGACACATTTCAAGTGAAGAAAGCTTTGATGACCTGGAAGTG 5979
QY 2131 PheAspGlyProSerGlyGlnSerProLeuLeuLysValaLeuSerGlyAsnTyrSerAla 2150
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 5980 TTTGATGTTCTTCTGGGCAAGTCTCTGCTAGAGCTTAAAGTGGAGATCATCTGAA 6039
QY 2151 ProLeuIleValThrSerSerSerAsnSerValTyrLeuAlaGlyTrpSerSerAspHisAla 2170
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6040 CAATCAAAATTTTACAGAGAGATGATGATTAATTCCTCGCTGATCCAGACCATGCC 6099
QY 2171 TyrAsnArgLysGlyPheLysIleArgTyrSerAlaProTyrCysSerLeuProArgAla 2190
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6100 ACCAGTAGAAGAGATTCAGATTCGATGACACACTTCTCAGATTTACACCAACCC 6159
QY 2191 ProLeuHisGlyPheIleLeuGlyGlnThrSerThrGlnProGlyGlySerIleHisPhe 2210
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6160 CTGAAGAAATGGGGTATTCTTAAACAGACTGCAGAGCGGTGGAAACAAAGTGCATTAT 6219
QY 2211 GlyCysAsnAlaGlyTyrArgLeuValGlyHisSerMetAlaIleCysThrArgHisPro 2230
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6220 TTTTGCAGCCTGATACCGAATGTGTGGCCACAGCAATGCACTGTAGCGAAACCA 6279
QY 2231 GlnGlyTyrHisLeuTrpSerGlnAlaIleProLeuCysGlnAlaLeuSerCysGlyLeu 2250
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6280 CTGGCATGTATCAGATGGGACTCCCTCAGCCACTGTGCAGAGCTGTGTCTGTGAATC 6339
QY 2251 ProGlnAlaProLysAsnGlnLysMetValPheGlyLysGlnTyrThrValGlyThrVala 2270
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6340 CCAAAATCCCGGAGAAAGGTTCAATTAACGGGAAACAGATTCCTTGGACAGTAAAGTG 6399
QY 2271 ValTyrSerCysSerGlnGlyTyrHisLeuGlnAlaGlyAlaGlnAlaThrAlaGlnCys 2290
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6400 GTCTATGATGTCTATGAGGGCTTCAAGCTTGAATCCAGCAGCAAGCAAGCAAGCGTGT 6459
QY 2291 LeuAspThrGlyLeuTrpSerAsnArgAsnValProProGlnCysValProValThrCys 2310
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6460 CAAGAAGATGGCTGTGAGACTTAAACAGGGGAAACCCGCCAGTGAAGCGGTGCTGC 6519
QY 2311 ProAspValSerSerIleSerValGlnHisGlyArgTyrArgGlnLeuIlePheGlnThrGln 2330
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 6520 CCAAGCATTAAGACTAGCTCTCAAGAACATGTATCTGAGGGCTGTGATTCAGGATCTTG 6579
QY 2331 TyrGlnPheGlnAlaGlnLeuMetLeuIleCysAspProGlyTyrTyrTyrThrGlnGln 2350

```



```
Db 6580 AATGAGTACGGTCTCAAGATTCTGAGCTGAGCTGCTTACTTACCTTAAAGAGCTGG 6639
      ::::: |||||::: ||| |||||
Qy 2351 ArgVal1LeaArgCysGln1LaenGlyLysTrpSerLeuGlyAspSerThrProThrCys 2370
      ::::: |||||::: ||| |||||
Db 6640 AGGCTCTCGCGTCCAGGCGCAATGGGACGTGGAACATGAGTGAAGCGCCAACTGCT 6699
      ::::: |||||::: ||| |||||
Qy 2371 ArgLe1LeaSerCysGlyLysLeuPro1LeaProbaenGlyAsaGlyLeuThrLeu 2390
      ::::: |||||::: ||| |||||
Db 6700 CAGATTACTGCTGTAAGCTTTCTTTCTCCCAATGGCAACAGATTGGAACCTG 6759
      ::::: |||||::: ||| |||||
Qy 2391 SerValTyrgLysAlaThrAla1LeaPheSerCysAsnSerGlyTyThrLeuValGlySer 2410
      ::::: |||||::: ||| |||||
Db 6760 ACGTTTATGGGGCCACAGCATATTATTCGTGCAACCGGCTTACAGCTTGGGGTCT 6819
      ::::: |||||::: ||| |||||
Qy 2411 ArgVal1ArgGlyCysMetAla1aenGlyLeuTrpSerGlySerGlyValArgCysLeuAla 2430
      ::::: |||||::: ||| |||||
Db 6820 CATGTCAGAGTCTGCGCAATGGGCTGTGAGCGGCGAGCGAACTCGATGCTGGCT 6879
      ::::: |||||::: ||| |||||
Qy 2431 GlyHicGlyGlyThrProGlyLysPro1Lea1aenGlyHic1LeaenGlyLysLeuThrSer 2450
      ::::: |||||::: ||| |||||
Db 6880 GGCACATCGCGTCTCCAGAGCCGATTTGAACTGACATGATGGAAGTGGCTTCACT 6939
      ::::: |||||::: ||| |||||
Qy 2451 TyraArgLysSerVal1Val1TyrgLysCysAsn1aGlyPheArgLeu1LeGlyMetSerVal 2470
      ::::: |||||::: ||| |||||
Db 6940 TACAGACACACGCTGTTACACTGACATCTCGTTCCGGCTTGGGAACTTCCGCT 6999
      ::::: |||||::: ||| |||||
Qy 2471 Arg1LeCysGlnGln1aenRhi1a1aTrpSerGlyLysThrProPheCysValPro1LeaThr 2490
      ::::: |||||::: ||| |||||
Db 7000 AGGTTATGCGCTGCACAAAGCCCAAGTGGTCTGCAAAAGCGCTGTGTCCTCCATCA 7059
      ::::: |||||::: ||| |||||
Qy 2491 CysGlyHicProGlyLysAsnProVal1aenGlyLeuThrGln1aenGln1PheAsnLeuAsn 2510
      ::::: |||||::: ||| |||||
Db 7060 TGTGTCACCTCGTGAACCTCGCCAGCATTTCACTAATGGCATGTGATCAACTAAT 7119
      ::::: |||||::: ||| |||||
Qy 2511 AspVal1Val1LysPheVal1CysAsnProGlyTyraMet1aGln1a1a1aArgSerGln 2530
      ::::: |||||::: ||| |||||
Db 7120 GATTCCTGTAATTCACCTGCAACACGGGCTATTGCTGCAAGGCGTGTCCGAGCCAG 7179
      ::::: |||||::: ||| |||||
Qy 2531 CysLeu1aSerGlyLysTrpSerAspMetLeuProThrCysArg1Le1LeaenCysThr 2550
      ::::: |||||::: ||| |||||
Db 7180 TGTGCGAGCAACGCGCCAGTGGAGTACCTCTGCCACGTGTCAGTGGTAACTGTTCT 7239
      ::::: |||||::: ||| |||||
Qy 2551 AspProGlyHicGln1aenSerVal1ArgGln1a1a1aSerGlyPro1a1aRhe 2570
      ::::: |||||::: ||| |||||
Db 7240 GATCCAGGCTTGTGAAAAGCCATTCGTACACGGGCAACAACTTCCCTGACAGTTT 7299
      ::::: |||||::: ||| |||||
Qy 2571 SerPheGlyThrThrVal1SerTyraArgCysAsn1aGlyPheTyraLeuLeuGlyThrPro 2590
      ::::: |||||::: ||| |||||
Db 7300 GAGTATGGAATGATATCTCTGACCATTCGCAAGAGGATTTTACTTGGGATCTTCA 7359
      ::::: |||||::: ||| |||||
Qy 2591 ValLeuSerCysGln1aenPglyTyThrTrpAspArgProArgProGln1CysLeuLeuVal 2610
      ::::: |||||::: ||| |||||
Db 7360 GCCTTGAACCTGTATGGCAATGGCTTATGGGACCGATCCCTGCCAAGTTTGGCTATA 7419
      ::::: |||||::: ||| |||||
Qy 2611 SerTyraGlyHicProGlySerProProHic1aSerGlnMetSerGlyAspSerTyThrVal 2630
      ::::: |||||::: ||| |||||
Db 7420 TCGGTGAGACACCCAGGGGCTCCGCAACGCTCTCACTGAGAGCTGTATTACCTAT 7479
      ::::: |||||::: ||| |||||
Qy 2631 Gly1a1aVal1a1aGlyTyraSerCys1LeGlyLysArgThrLeuVal1GlyAsnSerThrArg 2650
      ::::: |||||::: ||| |||||
Db 7480 GGGCCCTCTGCTGACTACTCTCTGCAAGGAGGAGCGACCTCATAGGCAACGACAGCA 7539
      ::::: |||||::: ||| |||||
Qy 2651 MetCysGlyLeuAspArgLys1aTrpThrGlySerLeuProHicCysSerGlyTyThrSerVal 2670
      ::::: |||||::: ||| |||||
Db 7540 GTGTGCGACGAAACAGTCACTGGAGCGGGGCACTGCCCTCACAGGAAATAATCT 7599
      ::::: |||||::: ||| |||||
Qy 2671 Gly1a1aCysGlyLysAspProGly1LeaPro1a1a1aGlyLeaGln1aenPheAsp 2690
      ::::: |||||::: ||| |||||
Db 7600 GGATTCTGTGTATCCGGGAGCCGACGACATGGGCTCGCTTGGTGGATGACTTTAAG 7659
      ::::: |||||::: ||| |||||
Qy 2691 ProGlyHicVal1aMetArgPheSerCysGln1aGlyHicVal1LeuAspGlySerSerGly 2710
      ::::: |||||::: ||| |||||
Db 7660 ACAAAAGTCTCTCCGCTTCTCTGTAATGGGACCCAGCTGAGGGCTCCCTGAA 7719
      ::::: |||||::: ||| |||||
Qy 2711 ArgThrCysGln1a1aenGlySerTrpSerGlySerGlyProGlyCysGlyVal1LeaSer 2730
      ::::: |||||::: ||| |||||
Db 7720 CGACGCTTGTCTCAATGGGTATGCTGAGCTGAGCTGAGCGGTGTGTGAGCGCGTCC 7779
      ::::: |||||::: ||| |||||
Qy 2731 CysGlyLysAsnProGlyThrProSerAsn1aArgVal1a1aPheSerAspGlyLeuValPhe 2750
      ::::: |||||::: ||| |||||
Db 7780 TGTGCAACCTCTGGACACCCACCAACGAAATGATTTCTAGTATGATGCACTTCTTCT 7839
      ::::: |||||::: ||| |||||
Qy 2751 SerSerSer1LeaValTyrgLysCysArgGlyGlyTyThrAlaThrGlyLeuLeuSerArg 2770
      ::::: |||||::: ||| |||||
Db 7840 TCCAGCTCGGTCAATCATGCTCTGTGGAGGCTACCAACGCTCAGGGCTCATGACCG 7899
      ::::: |||||::: ||| |||||
Qy 2771 HicCysSerVal1aenGlyThrTrpThrGlySerAspProGlyCysLeuVal1LeaenCys 2790
      ::::: |||||::: ||| |||||
Db 7900 CATTCACACGCAATGGGACCTTGACAGGACGCTCCGACATGCAATTAATGATGT 7959
      ::::: |||||::: ||| |||||
Qy 2791 GlyAspProGly1LeaPro1a1aenGlyLeuArgLeuGlyLysAspPheArgTyraAsnLys 2810
      ::::: |||||::: ||| |||||
Db 7960 GGGATTCAGGACACACTAGCAATGGCATCTCACTTGGGACCGACTTACCTTCAACAG 8019
      ::::: |||||::: ||| |||||
Qy 2811 ThrVal1ThrTyrgLysCysVal1ProGlyTyraMetMetGlySerHicArgVal1SerValLeu 2830
      ::::: |||||::: ||| |||||
Db 8020 ACTGTGACTATCAGTGTAACTCCAGGCTATGTGATGGAACAGTCACTCCGCACTATT 8079
      ::::: |||||::: ||| |||||
Qy 2831 SerCysThrTyraAspArgThrTrpAsnGlyTyThrProVal1CysVal1a1aLeuMetCys 2850
      ::::: |||||::: ||| |||||
Db 8080 CGCTGTACCAAAACGCGAGGTGAATCCAGCAAACTGTGTGCAAAAGCGGTGTGT 8139
      ::::: |||||::: ||| |||||
Qy 2851 LysProProProLeu1LeaPro1a1aenGlyLysVal1aGlySerAspPheMetTrpGlySer 2870
      ::::: |||||::: ||| |||||
Db 8140 CTTCAAGCCGCGCGCGGTGCAAGATGGAACAGTGAAGGAAATTTCCGCTGGGCTCC 8199
      ::::: |||||::: ||| |||||
Qy 2871 SerVal1ThrTyra1aCysLeuGln1aGlyTyrgLysLeuSerLeuPro1a1a1aPheThrCys 2890
      ::::: |||||::: ||| |||||
Db 8200 AGCATTAATTAACAGCTGCATGACAGGTTTACAGCTCTCTACCTCCGCACTCTCTGT 8259
      ::::: |||||::: ||| |||||
Qy 2891 GlyGlyAsnGlySerTrpThrGlyGlyLeuLeuProGln1CysPheProVal1PheCysGlyAsp 2910
      ::::: |||||::: ||| |||||
Db 8260 GAAAGTTCGCGGGTGTGGAAGAGAGATCCCAAGTGTCTCTGTGTCTCGGAGAC 8319
      ::::: |||||::: ||| |||||
Qy 2911 ProGlyVal1ProSerArgGlyArgGlyAspArgGlyLysSerTyraSerSerVal 2930
      ::::: |||||::: ||| |||||
Db 8320 CTTGGCATCTCCCGCAGAGGCACTTATGCGGAAAGTTTCACTTAAGTCCGAAGTC 8379
      ::::: |||||::: ||| |||||
Qy 2931 SerPheSerCysHicProProLeuVal1LeuVal1GlySerProArgArgPheCysGlnSer 2950
      ::::: |||||::: ||| |||||
Db 8380 TTCTTCAGTGCAAATCTCATTTATATCTGTGGATCTCCAGAAAGTCCGAAGCT 8439
      ::::: |||||::: ||| |||||
Qy 2951 AspGlyTyThrTrpSerGlyTyrgLysProSerCys1LeaAspProThrLeuThrCysAla 2970
      ::::: |||||::: ||| |||||
Db 8440 GACGGCAGCTGGAGCGGCATACCAACCACTGATGATCTCTCTATACACCTGCGCA 8499
      ::::: |||||::: ||| |||||
Qy 2971 AspProGlyVal1ProGlnPheGly1LeGlnAsnAsnSerGlnGlyTyrgLysVal1GlySer 2990
      ::::: |||||::: ||| |||||
Db 8500 GACCTGTGTAAGCCACCTTGAATACAAATAAGCTCCAGAGCTATGGAGTTGAAGC 8559
      ::::: |||||::: ||| |||||
Qy 2991 ThrVal1LeuPheArgCysGln1aGlyTyraLeuLeuGln1aSerThrThrArgThrCys 3010
      ::::: |||||::: ||| |||||
Db 8560 ACGTTTCTTTCAGGTGCAAAAAGGTACATATTAAGTTCCAGATCCGACCTGC 8619
      ::::: |||||::: ||| |||||
Qy 3011 LeuProAsnLeuThrTrpSerGlyTyThrProProAspCysVal1ProHicHicCysArgGln 3030
      ::::: |||||::: ||| |||||
Db 8620 CTGGCAATTTTAACTAGATGATGGATACAGACCGAATGTATACCTCATGTGCTCAGACAG 8679
      ::::: |||||::: ||| |||||
Qy 3031 ProGlyHicProThrHic1a1aenVal1Gly1aLeuAspLeuProSerMetGlyTyThr 3050
      ::::: |||||::: ||| |||||
Db 8680 CAGAAATCCCGGACACCGCGAATGTGAGGCCATGATTTCTTACTTCCGCTACACC 8739
      ::::: |||||::: ||| |||||
Qy 3051 Leu1Le-ThrPro1aArgArg1aSerProSerArgVal1a1aProSerThr1aPro1 3070
      ::::: |||||::: ||| |||||
Db 8740 TTAGTGTACACCTGCCATCCAGGCTTTTCTCTGCAAGGGGATCTGAGCACAGAACTGT 8799
      ::::: |||||::: ||| |||||
```

QY 3070 aAArgaMecAlaAlaGlyAlaLaserArg 3080
Db 8800 AAAGCAGACATGAATGACAGAGAAAGTCC 8830
RESULT 8
LOCUS AX374893 12525 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 3 from Patent WO0210199.
ACCESSION AX374893
VERSION AX374893.1 GI:19169788
KEYWORDS
SOURCE Rattus rattus (black rat)
ORGANISM Rattus rattus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Welch, A. A. and Elliott, G. S.
TITLE C3b/c4b complement receptor-like molecules and uses thereof
JOURNAL Patent: WO 0210199-A 3 07-FEB-2002;
Amgen Inc. (US)
FEATURES
source
1..12525
Location/Qualifiers
/organism="Rattus rattus"
/mol_type="unassigned DNA"
/db_xref="taxon:10117"
<1..9288
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD26836.1"
/db_xref="GI:19169788"
/translation="DAGKVDGTRSVLYLTGSSVDPDLIVSMNQMLHLQSDSISGP
GKAVYQIEKGGCGDPIPAVGRKTSGLHGDTLFECAAFELVGEVETICQNN
QMSGNKSCVFSCEFNFTASGILISNVEYEGNMMNCWMLISPSGRTILRPD
DVEQDFLAKDQDISDITVLTGFSNENPQAOLAXGKVLREPSDSHTTGKNI
LYTREGNECHDPIPVNGRRRGRPLLSGSVSHCDGKQESSEITCLIQDNN
VMSSTVRECAFGCGHLTASGVLTPEGMFGYKDSLNCCEWLEARHSHKTFPRF
QTEVNYDTLEVRDPTSSPLIGYHGTQAPFLISTGNMYLLFTDSSASVGLI
HYESVTLSDSCDPIPVNGHRSNFGIRSTVTSCDPTLSDDDPLICEKNHW
NHALPSCDALCGYIHGKSGTVLSPGPDYPSNLCTWTIEVSHKQGNMHPHL
ESSHDYLLITEDGSFSEPVARLIGSVLPHITAGLFGNPAOLRISDSISGNT
TFAEDLEPCDDDEPVASRRIGRQVGTTLAFTPCQRTLEAGTLTCLGGRVW
SAPLRVCACGASVKNRGTLLSPNSYDNNHCTYKTELKAGLHLARFOL
PBGDTLVKTVDPGKIPNYGYRIRDDGHFDTVLVLSNPNHLENSGSDTAQFOLTY
TSFVLVCEDEPGIIPNYGYRIRDDGHFDTVLVLSNPNHLENSGSDTAQFOLTY
KPMPSCVACGGLVHATSGRLSPGPAVDNNLCTWTIEADPKTKSLHRYVD
ETAHDLKLVWDGVDNIIILKEWSSGLPEPDISTNSLITQDSDPRTSKSGSFQF
STSIATCNDPMPONGTRYGDSREPDITTPCQDPCYOLQAGAKTTCQLNNRPFQ
DDPSCIACGGNLTGPAGVILSPNYRPPPEKEDCMWIKVNPVIALIRKFSME
PSYDLHIYKGRKNSNPLISFGSQAPETISSGNSLFLARSDASVGLSKFSEK
EKREACFDGNTNNGRTIGTDLKSTVYQDSGYKIVDSIECVAGQAKPSMD
RALPACAPCGQYSGEVVLSPNYPHYTAGQICIVISIVYKEVVGGAPEVQTA
LNDLAEFLDTPHQAARLLSLSGSHGELPLATNOILLRFSAGSARAHFHYV
AVRRTDTCSSVPEPRYGRIGSEPSAGSIVAFECNPGVILQSGALRACQVPA
QMDTIPSCVPCSGNFTORRGTILSPGPEYGNLNCWKILISBSGCIQVITSF
ATEOMNDLEIHHGEMTARLUSFSGTTPALNISTQCLHQSIDSVAAGFHL
EYTVGLAACQEPALPSNGIKIDRYMVDVLSPOEPTVQOGSHDSOMGVARW
NYPSPICIAICGGLTSMGVSILSGPSPSNNLCTWKILSPGIGHAIOPLNFT
EANDYLEIONGPYHSSPMGQSGDPLASLSTHETLIFEDHSGNROGFLSY
QAYELONCDPDPAFONGPMINSYVSGISPECYGVILGLHPIVLTQCHQDMWY
PPEPCAPCGYVTSONGTISRGPDPEDPILKCLMLVTVRPGHGVITNFTLQTEA
VNDYIVMGDPDONSPOLGVSQNTLLETAYSTNOVLKRFHSDSNGSFVNLHAF
OLKRCPPRPVADLLTEDEDEIDFVKYCHPGYTLISDGLTCLKSQLEFQGS
PTCEQCCPANEVRETSVILSPGPNVNSQTCVMSIKVDPNFTLLEFDEPOSE
KQPDALFVDSGSGOPLVLVSGNTEQSNFTSRNHLIYMSDHAHSKGFIRY
AAVYGLSTLTKNGVLTAKTAGLSKVYPCPKPGMIGSNATCRNPNPGVOWMS
MAPLCOAVSCGIPKAPGNGSFTGNETLILSKVTECNCGRTKDAQCAVACQEGIM
SNRKGPTCKVPKPSIEGOLSEHVLMLVSSSLNAYAGVILSPGTYLQGOQLLO
CANGTMMTEBDRPRCKVISCSSLSFPNGKIGTLTWGATLFTCWGTYLVGSHV
REGLANGMSSETRCLAGHSPDIPVNGHISGDFYRDIWVYQCNPGRLVGTAV
RCLQDHKWSGQTPVCVPTTCGHPNPAHLNGSEFVINDLVNFTCHGTGLOASR
AACRSNGWSSPLPICRVVNCSDPGEVNAVHGOONPESPSEYGTISWYCKKGFYL

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 12525
Score: 10720.00 Matches: 1843
Percent Similarity: 79.52% Conservative: 436
Best Local Similarity: 64.31% Mismatches: 543
Query Match: 63.11% Indels: 44
DB: 6 Gaps: 5

US-10-016-248-2 (1-3104) x AX374893 (1-12525)

QY 195 AaPGLyGlyAlaPGLyAaPGLyThrValLeuTyMetSerGlyAlaAsnAlaCysSer 214
Db 1 GAATGCCGGAGAGTGGGAGACCCAGACCGCTTGTGACGT----- 42
QY 215 AaPserProHlaThrProGlySerArgIleProGlu-----SerMetSerGlyAsp 231
Db 43 -----CTTACAGGCTCCAGTCTCCCTTACATCGAGGATAGCAATAG 90
QY 232 ILeTPaGlyGlnLysTrpThrValLeuGluIleCysArgAspIleSerSerAspAla 251
Db 91 AATGGCTCCAC-----CTGACGTACAGACAGACAGC 120
QY 252 ArgSerGlySerValArgLysSerProLysThrSerAsnAlaValGluLeuAlaPro 271
Db 121 -----ATTGGTTCCTCA 132
QY 272 Gly-----ThrGluIleGluGlnLysCysGlyAspProGlyIlePro 286
Db 133 GAATTTAAAGCTGTACCAAGAAATCGAAGAGAGAGGCTCGGACCTCGCATCCCA 192
QY 287 AlaTyrgIyaGagGluGlySerArgPheHisIscIyaAspThrLeuLysPheGluCys 306
Db 193 GCTTACGGGAGCGGACTGGCAGCAGCTTCTTGACACGGGACACGCTCACCTTGAAGTGC 252
QY 307 GlnProAlaPheGluLeuValGlyGlnLysAlaIleThrCysGlnLysAsnAsnGlnTrp 326
Db 253 CAGGACGCTTTTGAAGCTGTGAGAGAGAGAGAGTATTCGTGACAGAAACACCACTGG 312
QY 327 SerAlaLysLysProGlyCysValPheSerCysPhePheAsnPheThrSerProSerGly 346
Db 313 TCCGCAACAAGCCCAACCTGTGTCTTCACTTCTTCAACTTACAGGGCTCCTCGGG 372
QY 347 ValValLeuSerProAsnTrpProGluLysAspTyrgIyaAsnHisIleuHisCysValTrpLeu 366
Db 373 ATCATCTGTGGCCAACTATCTGAGGAATATGGAACAACATGAATGTGTGTGGTTG 422
QY 367 ILeuAlaLysProGluSerArgIleLeuAlaPheAsnAspIleAspValGluPro 386
Db 433 ATTATATCTGAGCCCGGAGCCGAGTTCACCTATCTCAATATTCATGATGAGGCTT 492
QY 387 GlnPheAspPheLeuValIleLysAspGlyAlaThrAlaGluAlaProValLeuGlyThr 406
Db 493 CAGTTTACTTCTTGGGCAAGATGATGGGATTTGACATCACTACAGTCCCGGAGCT 552
QY 407 PheSerGlyAsnGlnLeuProSerSerIleThrSerSerGlyHisIleValAlaArgLeuGlu 426
Db 553 TTCTCTGGCAATAGAGTGCCTGACAGCTGGCCGACGAGACATTAAGTACGCTGAG 612

```
Qy 427 PheGlnThrAepHieSerThrGlyLeuArgGlyPheAenIleThrPheThrThrPheArg 446
Db 613 TTTAGTCCGACTCACTCAACACGGGAGGTTCAACATCATATACACCACTTTGGT 672
Qy 447 HisAngIucYsProAapProGlyValProValAngIyluYsArgPheGlyAapSerIeu 466
Db 673 CAGAACGAGTGTCACTGACCTCGGATCCCTGTGAATGACCGGCTTTGAGACAGGTTT 732
Qy 467 GlnLeuGlySerSerIleSerPheLeuCYsAapGlyIlePheLeuGlyThrGlnGlySer 486
Db 733 CTGTGGAGAGTCTGTGTCTTCCACTGTATATAGTGGCTTTGAGACCTCAGAGTTCT 792
Qy 487 GlnThrIleThrCYsValIleuLeuGlyIleValValIlePheSerAAlaValIleuArg 506
Db 793 GAGCTCATCATGATCTTTCAGAGATGGAACGTGGCTGAGCTCTACTGTCTCTCC 852
Qy 507 CYsGlnAlaProCYsGlyGlyIleuLeuThrSerProSerGlyThrIleLeuSerProGly 526
Db 853 TGTGAAGCTCTTGTGGGCACTGTACAGCTTCTAGTGGGTCATATTACTCTCAGAA 912
Qy 527 TyrProGlyPheThrYsAapAlaLeuSerCYsAlaTrpValIleGlnAlaGlnProGly 546
Db 913 TGGCAGAGATATTCAAAGATCTTTAATGCGAATGGGTCAATTGAAGCCAAACAGGA 972
Qy 547 TyrProIleYsIleThrPheAapArgPheLeuThrGluValAanTyrAapThrIleuGlu 566
Db 973 CATTCATCAAAATTAACATTGACAGGTTCAGACAGAGATCAATTATGATCTCTGGA 1032
Qy 567 ValArgAapGlyArgThrYsSerAlaProLeuIleGlyValIleYsGlyThrGlnVal 586
Db 1033 GTCCGGAGTGGGCAACCAAGCTCATCCCACTGATTGGGAGTACATGACCAAGGCT 1092
Qy 587 ProGlnPheLeuIleSerThsAanTyrIleuThrIleuLeuPheSerThsAapIleYs 606
Db 1093 CCAAGTTCCTCATCAGACAGGAGAACTAATGTAACCTGTATTCACATCAGACAGC 1152
Qy 607 HisSerAapIleGlyPheGlnLeuArgThrGluThrIleThrIleuGlnSerAapHisCYs 626
Db 1153 CGGCTAGTGTGGCTTCTCTCATCATCATAGAGTGTGACTCTTTGAATCTGACTCTGT 1212
Qy 627 LeuAapProGlyIleProValAanGlyGlnArgHisGlyAaAapPheThrValGlyAla 646
Db 1213 CTGACCCGGGCACTCCCTGTAAATGTCATCGGCATGGCAGTAACCTTTGGATCAGATCT 1272
Qy 647 LeuValIleThrPheSerCYsAapSerGlyThrIleuSerAapGlyGluProLeuGlyCYs 666
Db 1273 ACACTGACCTTCAGCTGTGACCTCGGTACACGCTCATGTATGACATCCCTCATCTGT 1332
Qy 667 GluProAapPheGlnTrpSerArgAlaLeuProSerCYsGlnAlaLeuCYsGlyIlePhe 686
Db 1333 GAGAAAGAACATCACTGAGAACAGGCTTGGCCAGCTGTGATGCCCTGTGTGGAGGCTAC 1392
Qy 687 IleGlnGlySerSerGlyThrIleLeuSerProGlyPheProAapPheThrProAaAan 706
Db 1393 ATCATGAGAAAGAGTGGACTGTCTTTCACACAGATTTCCAGACTTTTATCCAAACTCT 1452
Qy 707 LeuAanCYsThrTrpIleIleGlnThrSerHisGlyYsGlyValPhePheThrPheHis 726
Db 1453 CTGAACGTATCATGACCATTTGAAGTCTCTCATGTGCAAGGAGGAGCATGAATTTTCCAC 1512
Qy 727 ThrPheHisLeuGlySerGlyHisAapThrIleuLeuIleThrGlnAanGlySerPheThr 746
Db 1513 ACCTTCACTTGAAGTTCCACGACTAATTGCTGATCAGACAGAGAGGAGTTTCTCA 1572
Qy 747 GlnProLeuArgGlnLeuThrGlySerArgLeuProAlaProIleSerAlaGlyLeuYs 766
Db 1573 GAGCGGAGAGCCAGGCTCATCGGTGCTGCTGCTCAGACACATTTAAGGCTGTGTTT 1632
Qy 767 GlyAanPheThrAlaGlnValArgPheIleSerAapPheSerMetSerTyrGlyIlePhe 786
Db 1633 GGAACATTCACCTGGGCAACTCAGGTTCTCATCTCTCATCTCTCATCTCTAAGAGGCTTC 1692
Qy 787 AsnIleThrPheSerGlyThrYsAapLeuGluProCYsGlyIleGluValProAlaTrp 806
1693 AACATTACGTTTGGAGATATGACCTGAACCCCTGTATGACCCCTGGAGTCCCTGCTAC 1752
Qy 807 SerIleArgIyluGlyLeuGlnPheGlyValGlyAapThrIleuThrPheSerCYsPhePro 826
Db 1753 AGTCGCGAATTTGGGTTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1812
Qy 827 GlyTyrArgLeuGlyIleThrAlaArgIleThrCYsIleuGlyGlyAaArgArgLeuTrp 846
Db 1813 GGATACCGCTTGAAGGTGCAACCAAGCTTACCTGCTGGGTGGGAGAGCGCCGAGTTGG 1872
Qy 847 SerSerProLeuProArgCYsValAlaGlyCYsGlyYsAanSerValIleThrGlyThrGlnGly 866
Db 1873 AGTCACACTCTGCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1932
Qy 867 ThrIleuSerProAanPheProValAanTyrAasAanAanHisGlyCYsIleYsSer 886
Db 1933 ACATTACTCTCTCCAAATTTCCCATCCAAATATATATATATATATATATATATATATAA 1992
Qy 887 IleGlnThrGlnProGlyYsGlyIleGlnLeuYsAlaArgAlaPheGluLeuSerGlu 906
Db 1993 ATGAAACAGAACCCGAAAGGAGATCATCTCAGACCCGAACTTTCACACTTTCGAA 2052
Qy 907 GlyAapValIleuYsValIleYsAapGlyAasAanAanSerAlaArgLeuLeuGlyValPhe 926
Db 2053 GAGACACACTTAAGTTATGTATGAAAGACAGCTCTCGAGTCACTGAGTGGAGTCTTC 2112
Qy 927 SerHisSerGlyMetMetGlyValIleThrLeuAanSerThsSerSerIleuTrpLeuAap 946
Db 2113 ACAAGAGTAACTAAGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2172
Qy 947 PheIleThrAapAlaGluAanThrSerIleGlyPheGluLeuHisPheSerSerPheGlu 966
Db 2173 TTCAACTTAACGGGTCAAGTACCGCCCAAGGCTTCAAGCTCACCTACACCACTGAGGCTGAG 2232
Qy 967 LeuIleYsCYsGlnAapProGlyThrProLeuAapIleYsValHisAapGlyIleYs 986
Db 2233 CTAGTGAATGTAGATTCAGAGATCCCTCACTATGCTCAAGATCCAGATGATGATGAT 2292
Qy 987 HisPheAlaGlySerSerValSerPheSerCYsAapProGlyYsSerLeuAapGlySer 1006
Db 2293 CACTTCAACAGACACTGTGGTCTCTCAAGCTGTCAACCCAGGCTACGCAATGATGAGGAGC 2352
Qy 1007 GluGluLeuLeuCYsIleuSerGlyIleArgArgThrTrpAapArgProLeuProThrCYs 1026
Db 2353 AGTACCTGACCTGCTCAAGTGTGGAGCCGAAAGGTGTGGACAAACCTATGCTTCTGT 2412
Qy 1027 ValAlaGlyCYsGlyIleThrValArgGlyIleValSerGlyGlnValIleuSerProGly 1046
Db 2413 GTGGCGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2472
Qy 1047 TyrProAlaProTyrGluHisAanLeuAanCYsIleTrpThrIleGlnAlaGlnAlaGly 1066
Db 2473 TACCTGCCCCATATGACCAACACTTATGCACTGTGACATTAAGAGGCTGATCGGCG 2532
Qy 1067 CYsThrIleGlyLeuHisPheLeuValPheAapThrGluGluValHisAapValLeuArg 1086
Db 2533 AAGACCAAYCAGCTCCATTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2592
Qy 1087 IleTrpAapGlyProValIleuSerGlyValIleuLeuYsGlyLeuSerGlyProAlaLeu 1106
Db 2593 GTCTGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2652
Qy 1107 ProYsAapLeuHisSerThrPheAanSerValValIleuGlnPheSerThrAapPhePhe 1126
Db 2653 CTGAGAGACATCCACAGACCTTCAACTGCTGACCTGAGTTCGATGAGACTTCTTC 2712
Qy 1127 ThsSerIyluGlnGlyPheAlaIleGlnPheSerValSerThrAlaThrSerCYsAanAap 1146
Db 2713 ATCAGCAAGTCCGGCTTCTCATCACTTCTCACTTCACTTCACTTCACTTCACTTCACTTCA 2772
Qy ProGlyIleProGlnAanGlySerArgSerGlyAapSerTrpGluAlaGlyAapSerThr 1166
1147 ProGlyIleProGlnAanGlySerArgSerGlyAapSerTrpGluAlaGlyAapSerThr 1166
```

```
Db 2773 CCTGGATGCTCAGAAATGGAACCCGCTATGCTGACAGCCGGGAACCTGAGAACCATC 2832
Qy 1167 ValPheGlnCyAspProGlyThrAlaLeuGlnGlySerAlaGluIleSerCyValLys 1186
Db 2833 ACCTTCAGTGTGACCTTGAGATACAGCTCCAAAGGCAAGCCAAATACCTTGCTGAG 2892
Qy 1187 IleGluAsnArgPhePheTyrGlnProSerProThrCysIleAlaProCysGlyVal 1206
Db 2893 CTTAACAACCGCTTCTTGCCACACAGCCCTCGTATGATGATGCTGCTGCTGCTGCTG 2952
Qy 1207 AspLeuThrGlyProSerGlyValIleLeuSerProAsnTyrProGluProTyrProPro 1226
Db 2953 AATCTGACAGCGCTGCTGAGATGATTTATCCCAAACTACCAACGCCATACCTCTCT 3012
Qy 1227 GlyGlyGluCyAspTyrPheValThrValSerProAspTyrValIleAlaLeuValPhe 1246
Db 3013 GGGAAAGAGTGTGATGAGAAATTAGGTGAACCCAGACTTTGTCATTTGCTTAAATATTC 3072
Qy 1247 AsnIlePheAsnLeuGluProGlyTyrAspPheLeuHisIleTyrAspGlyArgAspSer 1266
Db 3073 AAAAGTTTAGCATGAGACCAAGTTACGACTTCTCGCATATCTATGAAGGGAAGACTCC 3132
Qy 1267 LeuSerProLeuIleGlySerPheTyrGlySerGlnLeuProGlyArgIleGlySerSer 1286
Db 3133 AACAGCCACTGATGAGAGCTTCAGGGTTCAGAGCCCAAGAGAGATGAGACAGT 3192
Qy 1287 SerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPheVal 1306
Db 3193 GGTAACAGCTCTTCTCGGATTCAGAGAGTATCCCTCTGTGGCCCTGCGGGTTTGGC 3252
Qy 1307 IleAspTyrThrGluAsnProArgGlySerCyPheAspProGlySerIleLeuAsnGly 1326
Db 3253 AATGAATTTAAAGAAACCAACGGGAAGCTTGTACCTTGAGGAACATTAAGAACGGG 3312
Qy 1327 ThrArgValGlySerAspLeuLeuGlySerSerValThrTyrTyrCysHisGlyGly 1346
Db 3313 ACAAGATGAGAAAGAGACTTTAACTGGCTCTACAGTTACCTTACATGATGACTCTGGT 3372
Qy 1347 TyrGluValGluGlyThrSerThrLeuSerCysIleLeuGlyProAspGlyLysProVal 1366
Db 3373 TACAAAGATTGTGATCCCTCATCATGAGTGTGTGACAGGGGTGATGGGAAGCCGTCC 3432
Qy 1367 TyrAsnAspProArgProValCysThrAlaProCysGlyGlyGlnTyrValGlySerAsp 1386
Db 3433 TGGGACCGGGACATGCTGCTGCTGCAAGACACCTGTGAGAGCCAAATACAGGGCTCGAG 3492
Qy 1387 GlyValValLeuSerProAsnTyrProGlnAsnTyrThrSerGlyGlnIleCysLeuTyr 1406
Db 3493 GGGGTAGTTTGTCAACCAATACCTCATATACATACAGCGCTGGGAGATATGACTAT 3552
Qy 1407 PheValThrValProLysAspTyrValValPheGlyGlnPheAlaPhePheHisThrAla 1426
Db 3553 TCCATCCCGGTGCGCAAGAAATTTGGTGTGTGGACAGTTTCCATTTCCAGACTCGG 3612
Qy 1427 LeuAsnAspValValGluValHisAspGlyHisSerGlnHisSerArgLeuLeuSerSer 1446
Db 3613 CTGAGAGACTTGGCAGAAATTTGTTGATGAAACCATTCCTCAGCGCTCTCAGTTCT 3672
Qy 1447 LeuSerGlySerHisThrGlyGlnSerLeuProLeuAlaThrSerAsnGlnValLeuIle 1466
Db 3673 CTCTCTGCTTCCATTCAGGTGAACATCCCGCTGGTATCATTCATCATGATTCGCT 3732
Qy 1467 LysPheSerAlaLysGlyLeuAlaProAlaArgGlyPheHisPheValTyrGlnAlaVal 1486
Db 3733 CGCTTACAGCGCAAAAGCGGAGCTTGTGACCGGGGTTTCCATTCCTACCAACCCGTC 3792
Qy 1487 ProArgThrSerAlaThrGlnCysSerSerValProGluProArgTyrGlyLysArgLeu 1506
Db 3793 CCAAGCAGCAGTGAACAGAGTGCAGCTCCGTCCTGAGCCCAAGATATGGGAGAGGAT 3852
Qy 1507 GlySerAspPheSerValGlyAlaIleValAlaArgPheGluCyAsnSerGlyTyrAlaLeu 1526
Db 3853 GGTTCTGAGTTCTCTGACGGCTTCATCGTCCGATTCGAGTGCACACCGATTACTGCTG 3912
Qy 1527 GlnGlySerProGluIleGlyCysLeuProValProGlyValAlaLeuAlaGlnTyrAsnVal 1546
Db 3913 CAAGGCTCCACAGCCATCCGTTGTGCTGTGCGCAAAAGCTTTGGCCAGTGGATGAC 3972
Qy 1547 SerAlaProThrCysValValProCysGlyGlyValAsnLeuThrGluArgArgGlyThrIle 1566
Db 3973 ACCATCCCAAGCTGTATGATTCATGACAGTGGCAATTTCACTCAGAGAAAGAGGCACAA 4032
Qy 1567 LeuSerProGlyPheProGluProTyrLeuAsnSerLeuAsnCyValTyrLysIleVal 1586
Db 4033 TTATCTCCAGGCTACCTGATGAGCCCTATGGAACAACCTGAATCTGTATGAGAAATCA 4092
Qy 1587 ValProGluGlyAlaGlyIleGlnIleGlnValValSerPheValThrGluIleAsnTyr 1606
Db 4093 GTATCGAAGGCTCAGAGGATCCAGATCCAAATGATAGTTGCTTGCCACGAGAGAACTGG 4152
Qy 1607 AspSerLeuGluValPheAspGlyAlaAspAsnThrValThrMetLeuGlySerPheSer 1626
Db 4153 GACTCCCTGAGATCTCATGACGAGAGACATGACGGCCCCCAAGACTGGGAGCTTCTCA 4212
Qy 1627 GlyThrThrValProAlaLeuLeuAsnSerThrSerAsnGlnLeuTyrLeuHisPheTyr 1646
Db 4213 GGATACACAGTCCCGCACTGCTGAATGACCTCCAAACAGCTCTGCTGACATTCAG 4272
Qy 1647 SerAspIleSerValSerAlaAlaGlyPheHisLeuGluTyrTyrThrValGlyLeuSer 1666
Db 4273 TCGAATCATAGTTGTGCGCTGCGGCTTTTCACTCGGAATTCAAACCGTGGGTCTGCT 4332
Qy 1667 SerCyProGluProAlaValProSerAsnGlyValLysThrGlyGluArgTyrLeuVal 1686
Db 4333 GCGTCCAGGAACCTGCTCTCCAGGCAACCGCATCAAGATAGAGACCTATATGGGTG 4392
Qy 1687 AsnAspValValSerPheGlnCysGluProGlyTyrAlaLeuGlnGlyHisAlaHisIle 1706
Db 4393 AACGATGTGCTGCTTCCAGTGGGAGCTGGTATACCTTGACAGGGCGCTCACACATT 4452
Qy 1707 SerCysMetProGlyThrValArgArgThrAsnTyrProProProLeuCyIleAlaGln 1726
Db 4453 TCTTGTATGCGGGAACTGTAGTGGCGGAATATCTTCCCTGTGTGATTTGCCACC 4512
Qy 1727 CysGlyGlyThrValGluGluMetGluGlyValIleLeuSerProGlyPheProGlyAsn 1746
Db 4513 TGTGTGTGAGACATGACAGCATGATGAGATGATCTGAGGCCAGGCTTCCAGGCTCA 4572
Qy 1747 TyrProSerAsnMetAspCysSerTyrLysIleAlaLeuProValGlyPheGlyAlaHis 1766
Db 4573 TACCCCAACAACTGAGCTGACCTGGAAGATATCCCTGCCATGTGCTATGTGACAT 4632
Qy 1767 IleGlnPheLeuAsnPheSerThrGluProAsnHisAspTyrIleGluIleArgAsnGly 1786
Db 4633 ATCCAAATTCGATTTCTCACTGAAGCAACCATATACCTCGAGATTCGAAATGGC 4692
Qy 1787 ProTyrGluThrSerArgMetMetGlyArgPheSerGlySerGluLeuProSerSerLeu 1806
Db 4693 CCTTACCAAGTATGCCAAGATGAGGACAGTTACAGTGGCCCTGACCTGCGCTGACTG 4752
Qy 1807 LeuSerThrSerHisGlyThrThrValTyrPheHisSerSerHisSerGlnAsnArgPro 1826
Db 4753 CTGAGCACACACATGAACCTTCATCGCTTCTTATATGATCAACCTCAGAAACCGCA 4812
Qy 1827 GlyPheLysLeuGluTyrGlnAlaTyrGluLeuGlnGluCyProAspProGluProPhe 1846
Db 4813 GGATTTAACTCAGTTTCAAGCTTATGATTTACAGACTGCGGAGCCGAGACCCGCACTTC 4872
Qy 1847 AlaAsnGlyIleValArgGlyAlaGlyTyrAsnValGlyGlnSerValThrPheGluCys 1866
Db 4873 CAGATATGGTTTCAATGATCAATCCGATTCAGCGTGGCCAGTGCATTCATTTAGTGC 4932
Qy 1867 LeuProGlyTyrGlnLeuThrGlyHisProValLeuThrCysGlnHisGlyThrAsnArg 1886
Db 4933 TACCCGGGCTATCTTGTGAGCCACCTGTGCTACCTGCGCAAGATGAGACTGACAGG 4992
```

Oy	1887	AsnTrpAspHisProLeuLeuProLysCysGluValProCysGlyGlyAsnIleLeuThrSerSer	1906
Db	4993	AACTGGAACTAACCTTTCTCCACGGGTGACCTCCCTGTGGATTAAGTAACTACACAG	5052
Oy	1907	AsnGlyThrValTyrSerProGlyPheProSerProTyrSerSerSerGlnAspCysVal	1926
Db	5053	AATGGACACATTTATTTCCCTGGGTGTTCCACAGCATTCACATTTCTGAAGAGCTGCTG	5112
Oy	1927	TrpLeuIleThrValProIleGlyHisGlyValArgLeuAsnLeuSerLeuLeuGlnThr	1946
Db	5113	TGGCTGGTCATGTCCTCCACAGGACATGAGTGAATCAATCTTACCTTGCTGCAGACT	5172
Oy	1947	GluProSerSerGlyAspPheIleThrIleTrpAspGlyProGlnGlnThrAlaProArgLeu	1966
Db	5173	GAGCGCTGAATGACTATCATCGCTGTGTGGATGCTCTGCACGAATCTCCCTACGCTC	5232
Oy	1967	GlyValPheTrpAspSerMetAlaLeuIleThrValGlnSerSerSerAsnGlyValLeu	1986
Db	5233	GGGGCTTTCAGTGGAAACACTGCTCCCTGAGACAGATACACTCACCACACAGTCTTG	5292
Oy	1987	LeuIlePheHisArgAspAlaAlaThrGlyGlyIlePheAlaIleAlaPheSerAlaTyr	2006
Db	5293	CTCAAAATTCACAGAGATTTCTCCAAATGAGGCTTCTTTGTCCTCAATTTTCATGACATTT	5352
Oy	2007	ProLeuThrLysCysProProProThrIleLeuProAsnAlaGluValAlaThrGluAsn	2026
Db	5353	CAACTGGAAGAGTGGCCCGCTCTCCATAGATGGCCGAGGTGCATCTGCTTACAGAAAT	5412
Oy	2027	GluGluPheAsnIleGlyAspIleValArgTyrArgCysLeuProGlyPheThrLeuVal	2046
Db	5413	GAAAGCTTTGAATATGGGAGCTTGTGTAAAGTACAGTGCATTCAGGTATACAGCTGTGG	5472
Oy	2047	GlyAsnGluIleLeuThrCysLysLeuGlyThrTyrLeuGlnPheGluGlyProProPro	2066
Db	5473	GGAAGTGACACCTCGACATGCAAGCTCAGCTCACAGCATTTGTTCAAGCTCTCCACT	5532
Oy	2067	IleCysGluValHisCysProThrAsnGluLeuLeuThrAspSerThrGlyValIleLeu	2086
Db	5533	ACCTGTGAGCACAATGCCACGCCAATGAATGAGCAGCATCTTCTGGGGTATTTCTC	5592
Oy	2087	SerGlnSerTyrProGlySerTyrProGlnPheGlnThrCysSerTrpLeuValArgVal	2106
Db	5593	AGTCTGGGTACCCACGGCACACTATTTTAATCTCCAGACATGTGCTGGAGTATTAAGTG	5652
Oy	2107	GluProAspTyrAsnIleSerLeuThrValGlnTyrPheLeuSerGlyLysGlyTyrAsp	2126
Db	5653	GAGCCAAACTTAACTAATTAACGCTCTTTGTGGACACCTTCAAAAGTAAACAAATTTGAT	5712
Oy	2127	GluPheGluIlePheAspGlyProSerGlyGlnSerProLeuLeuValAlaLeuSerGly	2146
Db	5713	GCACGTGAGATTTGATGTTGTTCTTCTGGCAAAATGCTTTGTGTTAAGTGTAAAGTGGG	5772
Oy	2147	AsnTyrSerAlaProLeuIleValThrSerSerSerAsnSerValTyrLeuAspTrpSer	2166
Db	5773	AACCACTAGAACACTCCAAATTTTACACGACGAAGTAAACAATCTGATCTCCGTGGTCC	5832
Oy	2167	SerAspHisAlaTyrAsnAlaArgLysGlyPheLysIleArgTyrSerAlaProTyrCysSer	2186
Db	5833	ACAAGATCATGCACACGACAAAGAAAGATTCAGAATGCTGATGCACCTCTTACTGCAGC	5892
Oy	2187	LeuProArgAlaProLeuHisGlyPheIleLeuGlyGlnThrSerThrGlnProGlyGly	2206
Db	5893	CTCACTCTACACTCAAGATGATGTGGCTTTTAATAAATACCGAGGGCGCCCTGGGGAGC	5952
Oy	2207	SerIleHisPheGlyCysAsnAlaGlyTyrArgLeuValGlyHisSerMetAlaIleCys	2226
Db	5953	AAGGTGACGATTTTCTGCAGCCTGGATATGGAATGATGGCACACAGCAAGCCACCTGC	6012
Oy	2227	ThrArgHisProGlnGlyTyrHisLeuThrSerGluAlaIleProLeuCysGlyAlaLeu	2246
Db	6013	AGCGGAAACCAAGTGGGCGGTGTACCAAGTGGGACTGATGGACACGCTTTGCCAAGCTGTG	6072
Oy	2247	SerCysGlyLeuProGluAlaProLysAsnGlyMetValPheGlyLysGlyTyrThrVal	2266

Db	6073	TCCTGTGGAAATTC	CCAGAGGCTCCAGGAAATGGCTTCGTTCCAGGCAATGAACTTCA	6132
Oy	2267	G1YrThrYsAlaVal	TYrSerCYsSerGluG1YrThiSLeuGlnAlaG1YrAlaG1uaLa	2286
Db	6133	GACAGTAAAGGACTTA	TGAATGAATGAAGAGGCTTCAAGCTGGATGCGACATGACAAAGCC	6192
Oy	2287	ThiAlaGluCYsLeuAspThrG1YrLeuTrpSerAsnArgAsnVal	ProProGlnCYsVal	2306
Db	6193	ACTGCTGTGTCTCA	GAAGAATGGCTGTGGAGCAACAGAGAAAGCCACCCGCTGCMAA	6252
Oy	2307	ProVal	ThrCYsProAspValSerSerLLeSerValGlnHisG1YrAspArgLeuIle	2326
Db	6253	CCGGAGCCCTGGCC	CCAGCATCGAAGGCCAGCTGTCAAGACACGTCCTTGAGGCTGAT	6312
Oy	2327	PheGluThrArgGln	TrpGlnPheGlnAlaGlnLeuMetLeuIleCYsAspProG1YrTYr	2346
Db	6313	TCCGGATCAT	TGAATGAATATGAGAGCTCAAGTTCTCTCCAGCTGTACTGCTGCTACTTC	6372
Oy	2347	TYrThrG1YrGlnArgVal	ILeArgCYsGlnAlaAsnG1YrLysTrpSerLeuG1YrAspSer	2366
Db	6373	TTGCAGGGTCCAG	AGGCTGTGGCACTGTGCCAAGCCCAATGGACCTTGAAACACTGAGAGAGAC	6432
Oy	2367	ThrProThrCYsAspG1IleIleSerCYsG1YrGlnLeuProIleProProAsnG1YrHisArg	2386	
Db	6433	AGACCCAGATGA	ATAAGTCACTCTCTTGAGGAGCCTGTCTTCCCCCAATATGTAACAG	6492
Oy	2387	ILeG1YrThrLeuSerAla	TYrG1YrAlaThrAlaIlePheSerCYsAsnSerG1YrTYrThr	2406
Db	6493	ATAGGAGACGCTCA	TATATGATGAGCCACCCGCACTTTTACTCTGCATATACCGGTACACA	6552
Oy	2407	LeuValG1YrSerArgVal	ArgValCYsMetAlaAsnG1YrLeuTrpSerG1YrSerG1YrVal	2426
Db	6553	CTTTAGAGCTCC	CCATGTCCTGGAGTGGTGGCCATATGATGTCTTGAGACGGATCTGAACA	6612
Oy	2427	ArgCYsLeuAlaG1YrHisCYsG1YrThnProGlnProIleValAsnG1YrHisLeAsnG1Yr	2446	
Db	6613	AGGAGCTCGGGG	GTCACTGTGCTTCACAGACCCCATTTGAAATGGCCATATCAGTGGC	6672
Oy	2447	GluAsnTYrSerTYrArgG1YrSerValValTYrGlnCYsAsnAlaG1YrPheArgLeuIle	2466	
Db	6673	GATGGCTTCA	CGACTACAGGACACAGGTGTCTACCAATGCAACCTGGGTTTCCAGCTGTA	6732
Oy	2467	G1YrMetSerValArgIleCYsG1YrGlnAspHisHisIleTrpSerG1YrLysTrpProPheCYs	2486	
Db	6733	GGCAGCGTGTGAG	AGATTTGGCTGCACAGACCAACAAGTGTGTGGGGCAGACCCCGCTTGGC	6792
Oy	2487	ValProIleThrCYsG1YrHisIleProG1YrAsnProValAsnG1YrLeuThrGlnG1YrAsnGln	2506	
Db	6793	GTCCCACTTCAC	TGTGACACCCCTGGAAACCTTGCATGGCTCCACCAACCGCAGGAG	6852
Oy	2507	PheAsnLeuAsnAspValIleLysPheValCYsAsnProG1YrTYrMetAlaG1YrValA	2526	
Db	6853	TTCAACTGTAATGA	CTTTGGAATTTTCACTGCCATACGGGCTACTCTGCTGCAGAGGTGCC	6912
Oy	2527	AlaArgSerGlnCYsLeuAlaSerG1YrGlnTrpSerAspMetLeuProThrCYsArgIle	2546	
Db	6913	TCCGAGGCCAAT	TGCGAAGAACGGCCAGTGGAGGACCCCTTGCTTATCTGCGAGTG	6972
Oy	2547	ILeAsnCYsThrAspProG1YrHisGlnGluAsnSerValArgGlnValHisAlaSerG1Yr	2566	
Db	6973	GTGAACGTTC	CCGATCTCTGTGTGGAAATGCAATGTGCGCCACGGGCAACGAACTTT	7032
Oy	2567	ProHisArgPheAspPheG1YrThnThrValSerTYrArgCYsAsnHisG1YrPheTYrLeu	2586	
Db	7033	CCAGAAGATTTCC	AGTATGGAACAAGTGTATGATCACTGCAGAAAGGGGTCTTACTTA	7092
Oy	2587	LeuG1YrThrProValLeuSerCYsG1YrG1YrAspG1YrThnTrpAspArgProArgProGln	2606	
Db	7093	CTGGGCTCTTGC	CGCTGACCTGTGATGGCAAGTGGCTGTGGAGCGGCTCTTATCCGAG	7152
Oy	2607	CysLeuLeuValSerCYsG1YrHisIleProG1YrSerProProHisSerGlnMetSerG1YrAsp	2626	

Db 7153 TGTCTGCTATATCATATGAGGATCTCTGGGCTCCCGCTAATGCTGTCTCTGACTGAGAA 7212
 QY 2627 SerTyrThValGlyAlaValValArgTyrSerCysIleGlyValArgThrLeuValGly 2646
 Db 7213 TTGTTTACATTTGGAGCCACAGTTCACTCTCAAAAGGGGGCCAGATTCTCAACAGC 7272
 QY 2647 AasnSerThrArgMetCysGlyLeuAspGlyHisTrpThrGlySerLeuProHisCysSer 2666
 Db 7273 AATACACAAAGAGTCTGCCAAGAAAGACATCATGAGAGGATCCCTTCTCCCATTTTCA 7332
 QY 2667 GlyThrSerValGlyValCysGlyAspProGlyIleProAlaHisGlyIleArgLeuGly 2686
 Db 7333 GGAATATGCTCTGATTTTGTGTGATCCAGGACCCAGACATGGGTCTGCTTGGG 7392
 QY 2687 AspSerPheAspProGlyThrValMetArgPheSerCysGlyAlaGlyHisValLeuArg 2706
 Db 7393 GATGAGTTTAAACAAGAGTCTTTTGGATTTCTCTGTGATGGGACCACTCGCG 7452
 QY 2707 GlySerSerGluArgThrCysGlnAlaAsnGlySerTrpSerGlySerGlnProGluCys 2726
 Db 7453 GGTCTGCAGAGCCGACATGCTGGTGAATGGTCTCTGTCAGAGTCCAGCTGTGTGT 7512
 QY 2727 GlyValIleSerCysGlyAspProGlyThrProSerAlaArgValValPheSerAsp 2746
 Db 7513 GAGCGCTGCTCTGTGAACCTGGCACCCCTCAATGGAGATGATCTCAGACGAT 7572
 QY 2747 GlyLeuValPheSerSerIleValTyrGluCysArgGlyGlyTyrTyrAlaThrGly 2766
 Db 7573 GGATCCCTCTCTCAAGTCTGTATCTATCTCTGCTGGGAAGCTCAAGACCTCGGG 7632
 QY 2767 LeuLeuSerArgHisCysSerValAsnGlyThrTrpThrGlySerAspProGluCysLeu 2786
 Db 7633 CTCTATGAGCGCGACTGACAGCAAGCGGACATGACAGGACAGCCCTGACTGAC 7692
 QY 2787 ValIleAsnCysGlyAspProGlyIleProAlaAsnGlyLeuArgLeuGlyAsnAspPhe 2806
 Db 7693 ATCATGAGCTGTGTGATCTCTGGACACTGCGCCATGATGATCCAGTTTGGACAGCTTC 7752
 QY 2807 ArgTyrAsnValThrValThrTyrGlnCysValProGlyTyrMetMetGluSerHisArg 2826
 Db 7753 ACTTCAACAAGACCGTAGCTATCAGTGAACCTTGCTACTGATGAGCCCAACA 7812
 QY 2827 ValSerValLeuSerCysThrValAspArgThrTyrAsnGlyThrIleProValCysIle 2846
 Db 7813 TCACCAACCAATCCGCTGACCAAGATGATCATGATGATGACCGGCTCTGCAAA 7872
 QY 2847 AlaLeuMetCysIleProProleuIleProAlaAsnGlyValValValGlySerAspPhe 2866
 Db 7873 GCTGTTCTATGAGCCAGCTCTCTCTCACTGACCAAGCAAGGAGGAGGCTCAGCTTC 7932
 QY 2867 MetTrpGlySerSerValThrTyrAlaCysLeuGluGlyTyrGlnLeuSerLeuProAla 2886
 Db 7933 CGATGGGCTCCAGCAATACCTACAGTTGTGTGATGGCTACACAGCTCTCCACCTGGC 7992
 QY 2887 ValPheThrCysGluGlyAsnGlySerTrpThrGlyGlnLeuProGlnCysPheProVal 2906
 Db 7993 ATCTGTCTGTAAAGGCGCTGAGTATGAAAGAGAGTCTCTCAGTGTCTGCTGG 8052
 QY 2907 PheCysGlyAspProGlyValProSerArgGlyArgArgGlyAspArgGlyPheSerTyr 2926
 Db 8053 TTCTGTGGCGATCCAGGCACTCCAGACAGAGGAGCGCTCAGTGGGAAAGCTTACCTTT 8112
 QY 2927 ArgSerSerValSerPheSerCysHisAspProleuValLeuValGlySerProArgArg 2946
 Db 8113 AAGTCTGAGGTCTTCAATCCAGTGCAAAACCCCATTTGTGTAGTGGTCTCTCAGAGAGA 8172
 QY 2947 PheCysGlnSerAspGlyThrTrpSerGlyThrGlnProSerCysIleAspProThrIleu 2966
 Db 8173 ACCTGCAGGCGCATGGATGTGAGTGGCATCCAGCCCACTGTATGATCCACCCAC 8232
 QY 2967 ThrThrCysAlaAspProGlyValProGlnPheGlyIleGlnAsnAsnSerGlnIleTyr 2986
 Db 8233 ACCGCTTGGCCAGACCCCGGCACTCCCACTTGGATATCAGATATAGCTCGAAGAGATAC 8292

QY 2987 GlnValGlySerThrValLeuPheArgCysGlnIleGlyTyrLeuGlnGlnGlySerThr 3006
 Db 8293 GAGTTGGAAAGCATCTGTCTTCAATGATTAAGAAAGTTTACCATCCAGGCTCACT 8352
 QY 3007 ThrArgThrCysLeuProAsnLeuThrTrpSerGlyThrProProAspCysValProHis 3026
 Db 8353 ACCCGAGCTGTCTTCTCCAACTCACTGATGATGAAATCAGACAGAGTGCATCCCCAT 8412
 QY 3027 HisCysArgGlnProGlnThrProThrHisAlaAsnValGlyAlaLeuAspLeuProSer 3046
 Db 8413 GCTGCGCGGACGACCAAGACCCAGCGCATGAGATGAGAGCATGATCTTCAGCT 8472
 QY 3047 MetGlyTyrThrLeuIle 3052
 Db 8473 TTGGCTACACCTTAGTC 8490
 RESULT 9
 AB114605
 LOCUS
 DEFINITION
 Homo sapiens mRNA for CSMD3 protein isoform 2, transcript variant b, complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 Shimizu, A., Asakawa, S., Saeki, T., Yamazaki, S., Yamagata, H., Kudoh, J., Minoshima, S., Kondo, I. and Shimizu, N.
 A novel giant gene CSMD3 encoding a protein with CUB and sushi multiple domains: a candidate gene for benign adult familial myoclonic epilepsy on human chromosome 8q23.3-q24.1
 JOURNAL
 MEDLINE
 PUBMED
 12943675
 2 (bases 1 to 13040)
 Shimizu, N., Asakawa, S. and Shimizu, A.
 Direct Submission
 Submitted (11-JUL-2003) Nobuyoshi Shimizu, Keio University School of Medicine, Molecular Biology, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan (E-mail: nshimizu@med.keio.ac.jp, Tel:81-3-3351-2370, Fax:81-3-3351-2370)
 FEATURES
 source
 Location/Qualifiers
 1..13040
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8q23.3"
 /class="testis"
 1..13040
 /gene="CSMD3"
 172..11175
 /gene="CSMD3"
 /codon_start=1
 /product="CSMD3 protein isoform 2"
 /protein_id="BAC82444.1"
 /db_xref="GI:34330133"
 /translation="MMSWFLCQKPVOLDROTASGFYTCGGTLKGLNGTIESPFPYGYPNAGNTWVIAEERNRIQIVQSPALAEEDYVLSLYGHPHTPRRTGFRHP PMSSTKSPSLRLTSPDAVSAHQFKYVELOSSECGNVPKVLKTRFDVGGKI RYSCVAVGILIDGHPDITCIANSVNTASMPPEPICAEDACGATMGSSGIISSPFP NETHNADCTWTVIAEPGDTLSLFTDFQMEKRYDLLEIGSEPPRIWISGNNIPPI ISKNNMLRLHFVTDLSNRYRGSAPYQGSSTLHTSTGLSEHNRTTGAALAVASP ADVTVGSVAIVTHIRLSEORQVOTSLRNSGLDPTSKDGLSPHADTOTSTRRPPA EOJERKEIAVTVHRYKKAIDFKSRGFLPGKDNKRSIINEGIKTASNLCDPDC EPENGRIGSDPSIGSTVQSCDEDEVVLCAGASITQORLAEPFAMSDIRPCKVATC GSNLQGPSGCTFTSPNPFPQYDSNAOCVWITTAIVNTKVIQINFEEDLIGDTLTIG DGEEDVDPRIVQLVTSFVFDLIVSMSSQMLHQTDSVSVGKVAVKELKESK GDFGTPLVGIREDGFSNRDVLRFECQGFELIGKSIIVCOENQMSANIPIICIFPL

SNFPAWNGSIVPEYPCGNCTNSHILCWIISDPSGRHLSFNDPLSOPCPLAVXO
DSPEPLRGLTGTAEVPEYPCGNCTNSHILCWIISDPSGRHLSFNDPLSOPCPLAVXO
IPIAARFGDPLQJGSSISYICBEGIKTQGBLITCLIMDRKWMSSILPRCACPCCG
GHGASBGVJLSDMPQYVPEYPCGNCTNSHILCWIISDPSGRHLSFNDPLSOPCPLAVXO
PMLPSGLSNGVOYQVPLFSSNSIYLLFTDMSRNSNGKTHIYESVTNVYSCD
GJFVHGRAGHDPESISCTVSPSCDSYJSHHEPLRCKKNHMSHPLTCDLACGDG
VRSQSGIISRGVPEYRNLSNCTMYDTHCKGQUNPHHPLTJEDHLLTITENS
FTQPLARLSDPLPTIMAGVYRPAQRLFTSDSIYEBENITTESGTYNLEPCDDG
PQYQSGIRGNPQJGDTLPTSCSSGRALGSGEINILCLQGRARWASPLPRVCACAS
ATNNEGILISPNYPLANNHETCSIVQJQAKGKINSARTHTLQAGVJLIXGKOK
TLLHAGFASMRGLTSSNSQMLENSNDPEGEDFOLVYFSLSCBEGJF
OFKXISDOCHPAGTITVSCNGTQJLHSSILKCMTEBRAMOVPLSICABEGFR
KBSGSGILSPYRPPDNNLRKMMHIBDYGNISYLOPLAFDTASHDILRWDEPP
ENDMLKEIGSSILPESHTINLITVITPDTQYISKSQFALQESSYASACRQVPL
MNGRNDGDEHGGVTLQFQCDPEYLOGBERTICJOEGRYKWPQSPVCLAPCQVPL
TSSGLSILPNFPHYSHSDCMTITVANADYSLAFSIEBENDYFITYGDPSSIM
PLSIFGRSOKHPIPERLESSNMMHLAPRDSQVSTGHEILXALRESCDFDGNIS
NCTRLMDYKSGSYVITYCDAGVYLOGYSTLTCLIMGDRBERNARLPSCHACGSSS
TGSQJVLASHPYRNLYVGNQVSIYSLAVEPEYVQGFYFOTSLADHVYDQPTQV
SSLSLSLASHSGEYSLPESAGNOITRTFSVPLTAKGHFVYAPVPTSSQSDV
EPRGRKINIEFPAVASSVLEDCNPLJLIGSLALICEVYPLNADMSNLPTCPCG
GILKRGKGTILSPREYPPYNNLNCWKLTVEBAGIIOVYVSPATBNMSLPTFGYQ
GDNNAPLVSQGTITPHLNTGSNLLXNPDOSIYSVAGHLEYTLATJLSCDEPQ
TPSSGKITGRNMGVYVSCQJQSLGSHSITTCOMPEPRANRYPLTCLACGGGA
MSDSGVALSPREPGANVSSLDCTWTLINPLIGEWALQVFNSTETHDYLEVSSGS
ENGVINLSPQJPLSSLFETHTESILYHSPYSQNKQFHLVYAOLOSCDPREPP
RATVITINDLFSQVQJIEFCEPEYTLIGNSALTLCHGYSRNMNHELPELCAJGNT
AMNTQISPCQPEYQCOMPLVAPVPGNGIYNFTLOEYRHYDTITWADSDPP
SPQJGFSRSTALESYVSTNOLIKRSHDPTTSGEPYLSHAYLOVQCPYRPPVPA
BILTEDEFEIGDIRYQCLPGLTVGNALITCRJGEBRQMDGAPVQVCLCANELR
LDSTGVLSRQYSPYENLQOMCMSISYKRGNIIMPFEPRTYEPYVYDQYVPGNI
OSPVLSLQSGYSAPNTTNGHEVLELQMSADHNKNGFRTYALFCSPEYDQYV
YIISOTGGQVSNVYRACDGPPLTVEKSSAVKRSYGHAMDVAPKALOSIGIKR
APNNGELITLDYVGRVYRQNDYRSLSKELTAYVQOSDTNHNKPRQVYV
PSINPLIHEHRKRIYVNGSHYKTYVSCDPCDGHGLGAPASIBCLPNTGWSNREPP
YCOJISCGEPLTPPNKQJITQYVSYSTALPTCDJLPMVLGSAVRECLSSIMSEST
RCLAGHGOJPELNGVQJIEYENQYSDYVTCQDNBFRPLIGSVYICODHNNSGQJPL
RCVSPASGCHSPSLYGTGTSNGENGNVYVTFSCNIGMOGTQACQANOMHPP
CKVYVNSDPIGJANVRSREKIEHNGFTYGVYDNCNGYLFSSVLTICQPMGDG
RPEJECIMIDGHPVAPNVASLEGKTYTGSVYHISCTGRKSLQOSRTQJOLGHHMS
GQJQSGGADATGCGDGTGHSRQESRNSRTKSYVYRACDGYILSHESBERTLANG
SMTROJECACATGCGDGTGHSRQESRNSRTKSYVYRACDGYILSHESBERTLANG
GATGQJPLNCTIISCGBPJIPANGLYGDDYVQVGNVSMQCPGYEMLNSRIJRTCT
INGWMSVMPYCAVYVTCPPRQJONSLRLEGNPMDGSSISYICBEGIKTQGBLIT
GNGTWSGVRQCLPKFGCDGJLTAQGRKBEKSFITYOSESRECNFPLIVSSSTRICQ
ADGWSGSSIPESICPYTQJSCENGEYRPHASQUNNTGFQYVGVCHCKKHLQGSTIT
LDGJLSDPLSICPEICPHSCQKOPFPAHANVGNMLPBHGHTLYTQJOPFLAGGT
EHKRGADNTMTQJPLCEBGSKILYKNDPRALPGRSPSLVDPDPAQONTIKVSNV
FKGKORMTITVYNSFNASTORVNAVTLNSNMELLSGYKSOEALMARIYLIYPAH
ASVYKQKENWANDGFIASAPDQATVYVPOGFTQGDYQGFQJLQJLGNMSSGNSHQ
PHGNSNVSVALIARVPFALLPAGFEVYIKQPLTAKTQJTGCSYHNNNSQALFENP
YVINDNASVEKARLVPDNPNTVNTCYTV

Alignment Scores:					
Pred. No.:	0	Length:	13040		
Score:	10591.50	Matches:	1816		
Percent Similarity:	76.65%	Conservative:	501		
Best Local Similarity:	60.07%	Mismatch:	665		
Query Match:	62.36%	Indels:	41		
DB:	9	Gaps:	6		
US-10-016-248-2 (1-3104) x AB114605 (1-13040)					
OY	36	VallVylvlglnllleGlnLeulysSerArggIyVallylSleuMetProSeRlyLshspn	55		
Db	1393	GTGAAGAAAGCCATGATTATAAATCTAGAGCATTTTAATTGTTCCAGGAAAGACAAAC	1452		
OY	56	SergInlystHsrSerValeuthrGlnValGlsValserGlnGlytHtAsmMcyapRo	75		
Db	1453	AGCAACAAGTTTTCTACTTAAATGAGGAGAGTATATAAACAGCTCCAATTATGCCCA	1512		
OY	76	AaspProglIyleProgluArgGlyLyshArgleuglySerAaspheargLeuglySerseR	95		

Db	1513	GATCCAGAGAACCCAGAAATATGGGAAGAAGATCGCATCAAGTTTAAAGCTTTGATCAACT	1572:::
Oy	96	ValGlnPheThrCySaengluGlyTyraSpIeugInGlySerLySaRglIeThCyemE	115	
Db	1573	GTGCAAGTTCTCTGTGATGAAGATTATATGCTCTACAGAGCCGCAAGACATCACTCTGTCAA	1632	
Oy	116	LySaValSerAspMetPheAlaAlaITrSerAspHisArgProValCySaRglAlaArgMet	135	
Db	1633	CGAGTACTCGAAGTTTGTGCTGTGATGAGATACAGGCTGTGTGTAAGAGTAAACG	1692:::
Oy	136	CysAspAlaHisLeuArxGlyProSerGlyLleIleThSerProAsnPheProIleGln	155	
Db	1693	TGTGGCTCTAATCTTCAAGAGCCAAAGTGATACCTTTACATCTCCCAATCTTCGTTCCAG	1752:::
Oy	156	TyrAspAsnAsnAlaHisCySaValITrpIleIleThraIalaAsnProSerLySaValIle	175:::
Db	1753	TATGACAGCAATGACCAATGTGTGGGTGATCAGACAGAGTAATCAAAATPAAGTTATC	1812:::
Oy	176	LySaLeuAlaPheGluGluPheAspLeuGluArgGlyTyraPheITrLeuThraValGlyAsp	195	
Db	1813	CAGATPAATTTTGAAGAATTTGATCTGGAGATTGGCATATGATATACCTTGACATGGCGAT	1872:::
Oy	196	GlyGlyGlnAspGlyAsaPbGlnLySaThValLeuITyrMetSerGlnAsnAlaCySaSerAsp	215:::
Db	1873	GGGGGCGAAGTTGAGATCCTTAGACAGTGTCTTCAAGTG	1911:::
Oy	216	SerProHisITrProGlySerArgIleProGlu-----SerMetSerGlyAspIle	232	
Db	1912	-----CGACATGGAGCTTTGTATCCAGACCTTGATAGTACAGACATGATAGGCAATG	1962:::
Oy	233	TrpArgGlnLySaITrPheITrValLeuGluIleCySaArgAspIleSerSerSerAspAlaArg	252	
Db	1963	TGGCTGCACCTTCAACAG-----GACGAAAGT	1989:::
Oy	253	SerGlySerValaArgLySerProLyThSerAsnAlaValGluLeuValAlaProGly	272	
Db	1990	GTTGATCTGTGGTTTCAGGTTAACTACAA-----	2022:::
Oy	273	ThrGluIleGluGlnGlySerCyGlyAsaPProGlyIleProAlaITyGlyArgArgGlu	292	
Db	2023	---GAATTGAGAAAGAAAGTTGGTGATCTCTGTACACCTTATATGAAATTAGAGAA	2079:::
Oy	293	GlySerArgPheHisArgLyAspITrLeuLySaPheGluCyGlnProAlaPheGluLeu	312	
Db	2080	GGCGATGGATTTTCTAATCGATGTTTAAAGTTTGAATCCAGTTTGGGTTGAATT	2139:::
Oy	313	ValGlyGlnLySaAlaIleThCyGlnGlnLySaAsnAsnGlnITrSerAlaLyLyProGly	332	
Db	2140	ATTGAGAGAAATCCATGTTGTTGCAAGAGAAATPACAAATGCTGCAACATACCCATC	2199:::
Oy	333	CysValPheSerCyPhePheAsnPheThSerProSerGlyValaValLeuSerProAsn	352	
Db	2200	TGTATCTTCTCCCTGCTGTCTAATCTTATCTGACCAATGGGAACAGTTCTTCTCTGAT	2258:::
Oy	353	TyrProGluAspITyGlyAsnHisLeuHisCySaValITrPheIleLeuAlaArgProGlu	372	
Db	2260	TACCAGAGAGGTATGGAATTAATTTAAATGATCGACATGACGATTAATCTGATCGAGGG	2319:::
Oy	373	SerArgIleHisLeuAlaPheAsnAspITLeaSpValGluProGlnPheAspPheLeuVal	392	
Db	2320	AGCGGATACATCTTCTTCAAGACTCTTGACCTGGAAATCCAGTTTGATTTCCCTGCT	2379:::
Oy	393	IleLySaAspGlyAlaITrAlaGluAlaProValLeuGlyITrPheSerGlyAsnGlnLeu	412	
Db	2380	GTTAAAGTGTGACTCTCCAGAAATCCCAATCTTGGAACTTTACTGGGGCTGAGAGTG	2439:::
Oy	413	ProSerSerIleITrSerSerGlyHisLeuValaAlaArgLeuGluPheGlnITrAspHisSer	432	
Db	2440	CTTTCCATCTTACTAGTAATAGCACATACCTGCGATTGGAAATTTCAGGCTGACCATCA	2489:::
Oy	433	ThrGlyLySaArgGlyPheAsnIleThPheITrPheArgHisGlnGluCyProAsp	452:::

```
Dh 2500 ATGTCAGACGCTTGCTTAACATCACTTACCAACATTTGACATATGAAATGCCCTGAT 2559
Qy 453 ProGlyValProValAsnGlyIysArgPheGlyAspSerLeuGlnLeuGlySerSerIle 472
Dh 2560 CCTGGAAATACCAATCAATGACGCGGTTTGGGACAACCTTCAATATGAGAAATTCAT 2619
Qy 473 SerPheLeuCyAspGlnGlyPheLeuGlyThrGlnGlySerGlnThrIleThrCysVal 492
Dh 2620 TCAGTATTATTTGGAGAAAGAGATTATTAACCCAGGAGAACGAAACAAATTTACATGAT 2679
Qy 493 LeuValGlnGlySerValValIlePheAsnSerAlaValLeuArgCysGlnValIleProCysGly 512
Dh 2680 CTATATGAGTGAAGAAATATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2739
Qy 513 GlnIleLeuThrSerProSerGlyThrIleLeuSerProGlyIleProGlyPheThrIle 532
Dh 2740 GGCCATTTTTCAGCTCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2799
Qy 533 AspAlaLeuSerCysAlaIleTyrValIleGlnValIleGlnValIleProGlyIleThr 552
Dh 2800 GACTCTTTGAATTTGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2859
Qy 553 PheAspArgPheLeuThrGlnValAsnIleTyrAspThrLeuGlnValIleAspGlyArgThr 572
Dh 2860 TTGAAAGATTTCACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2919
Qy 573 TyrSerAlaProLeuIleGlyValIleTyrIleGlyThrGlnValIleProGlnPheLeuIleSer 592
Dh 2920 CTCTGTCACCCCTTGCTTGATCTTACATATGACCCCAAGTGGCCCACTTCTATTTAGT 2979
Qy 593 ThrSerAsnIleTyrLeuIlePheSerThrAspIleSerIleSerIleSerIleSerIle 612
Dh 2980 AGCAGTATTTTAAATACCTTCTATTTTAAACAAGACAGTGGTCCAAATATGATGCTTC 3039
Qy 613 GlnLeuArgIleGlnIleThrLeuGlnSerAspIleSerIleSerIleSerIleSerIle 632
Dh 3040 AAGATTCATTTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3099
Qy 633 ValAsnGlyGlnArgIleGlyAsnAspPheTyrValGlyValIleValIleThrPheSerCys 652
Dh 3100 GTATATGCGCGCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3159
Qy 653 AspSerGlyTyrThrLeuSerAspGlyGlnProLeuGlnCysGlnProAsnPheGlnIle 672
Dh 3160 GATTCAGAGATACAGGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3219
Qy 673 SerArgIleLeuProSerCysGlnValIleLeuGlyGlyPheIleGlnGlySerSerGly 692
Dh 3220 AGCATCACTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3279
Qy 693 ThrIleLeuSerProGlyPheProAsnPheTyrProAsnAsnLeuAsnCysThrTyrIle 712
Dh 3280 ACATCTTATCACTGCTGATACCGGAATTTTATCCAAATCTCTGAATTTGATCATGAG 3339
Qy 713 IleGlnThrSerIleGlyIleGlyValIlePhePheThrPheIleThrPheIleLeuGln 732
Dh 3340 GTTGAATGTAACCAATGAGAAAGTGTGAGTTCACTTCACTTCACTTCACTTCACTT 3399
Qy 733 GlnIleAspTyrIleLeuIleThrGlnValIleGlySerPheThrGlnProLeuArgGln 752
Dh 3400 CATCATGACTACTTACTGATCAAGAGATGAGATTTTATCCCAACCACTGACGCTG 3459
Qy 753 ThrGlySerArgLeuProAlaProIleSerAlaGlyLeuTyrGlyAsnPheThrAlaGln 772
Dh 3460 ACTGGTTTCAGATTTCTCCAAACAAATGCTGCTCTATGAAATTTTCAAGGCTCA 3519
Qy 773 ValArgPheIleSerAspPheSerMetSerTyrGlnGlyPheAsnIleThrPheSerGln 792
Dh 3520 TTGCGTTTCATTTCAATTTTCAATATGAAAGATTTTAAACAATCAATCTTCTGTA 3579
Qy 793 TyrAspLeuGlnProCysGlnGlnProGlnValIleProAlaTyrSerIleArgIleGlyLeu 812
Dh 3580 TATAACCTTGAACCTTGAAGATCTGCAATTCCTCAATATGATGATGATGATGATGAT 3639
Qy 813 GlnPheGlyValGlyAspThrLeuThrPheSerCysPheProGlyTyrArgLeuGlnGly 832
Dh 3640 AACTTTGGGATTTGTGACACTGACCTTCTATGCTCTTGGGTTATGATGATGATGATG 3699
Qy 833 ThrAlaArgIleThrCysLeuGlyGlyArgArgLeuTyrPheSerProLeuProArg 852
Dh 3700 ACATCAAGATCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3759
Qy 853 CysValAlaGlyCysGlyAsnSerValThrGlyThrGlnGlyThrLeuLeuSerProAsn 872
Dh 3760 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3819
Qy 873 PheProValAsnTyrAsnAsnAsnIleGlyCysIleTyrSerIleGlnThrGlnProGly 892
Dh 3820 TATCCACTCACTAATGAAACCAACCATGATGATGATGATGATGATGATGATGATGAT 3879
Qy 893 LysGlyIleGlnLeuValAlaArgAlaPheGlnLeuSerGlnGlyAspValIleVal 912
Dh 3880 AAGGAGATCAATATTTCAAGCCAGAACATTTCAATTAGACAAAGAGATGTTCTTAAGAT 3939
Qy 913 TyrAspGlyAsnAsnAsnSerAlaArgLeuLeuGlyValIlePheSerIleSerGlnMet 932
Dh 3940 TATGATGAAAGATTAAGAAACGATCACTACTAGTGTCTTACTGTGATCTATAGGCC 3999
Qy 933 GlyValIleThrLeuAsnSerThrSerSerLeuThrIleAspPheIleThrAspAlaGly 952
Dh 4000 GGACTGACACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4059
Qy 953 AsnThrSerLysGlyPheGlnLeuIleIlePheSerSerPheGlnLeuIleLysCysGln 972
Dh 4060 GGGACAGATGAAGCTTTCACTGTGTATACAGGTTTGAACCTGCACACTGAGAAAT 4119
Qy 973 ProGlyIleProLysPheGlyTyrIleValIleIleAspArgIleGlyIlePheAlaGly 992
Dh 4120 CTTGCACTTCCAACTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 4179
Qy 993 ValSerPheSerCysAspProGlyTyrSerLeuArgIleSerGlnLeuLeuLeu 1012
Dh 4180 ATCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4239
Qy 1013 SerGlyGlnArgGlyThrTyrAspArgProLeuProThrCysValAlaGlnCysGly 1032
Dh 4240 AAGGAGAGAAAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4299
Qy 1033 ThrValArgGlyGlnValIleSerGlyGlnValLeuSerProGlyTyrProAlaProTyrGln 1052
Dh 4300 CGTTTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4359
Qy 1053 HisAsnLeuAsnCysIleTyrThrIleGlnValAlaGlyCysThrIleGlyLeuHis 1072
Dh 4360 AATTAACCTGCTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4419
Qy 1073 PheLeuValIlePheAspThrGlnGlnValIleAspValLeuArgIleTyrAspGlyProVal 1092
Dh 4420 TTTCTGCTTTTGTATGAGAAAGATCAATGATATCTCCGATCTGGAGCGGTCCACA 4479
Qy 1093 GlnSerGlyValIleLeuLeuGlyLeuSerGlyProAlaLeuProLysAspLeuHisSer 1112
Dh 4480 GAAATATATGCTTTTAAAGAAATTAAGTGAATCTTATTCCTGAAGAAATTCATAGC 4539
Qy 1113 ThrPheAsnSerValIleGlnPheSerThrAspPhePheThrIleGlySerGlnGlyPhe 1132
Dh 4540 ACCCTCAATATATGATTAACCACTGATGATGATGATGATGATGATGATGATGATGAT 4599
Qy 1133 AlaIleGlnPheSerValSerThrAlaIleThrSerCysAsnAspProGlyIleProGlnAsn 1152
Dh 4600 GCAATTCAGTTTCAAGTCTGTTGCCACTGCGTGTGCGAGCCAGGGGTCCCAAGAT 4659
Qy 1153 GlySerArgSerGlyAspSerTyrGlnAlaGlyAspSerThrValPheGlnCysAspPro 1172
Dh 4660 GGAATCTGAAGATGGAGATGAAGAACTGGGAGACATGTTGTTTTCATATGATGACCA 4719
```


Oy	1173	GLYLYRALALEUGINGLSERIALAGLULESERCYVALVLSIEGLUASAPRAPH	1192
Db	4720	GGATATGAACTTCAAGAGAGAGAAAGAAATTAACCTGCATTACAGTAAAGAAATCGGACTTC	4779
Oy	1193	TRGILNPROSERPROCTHRCYSIELLEAPROCSGLIGLYASPLEUTHGILPROSER	1212
Db	4780	TGGCAGGCCCAACCCACAGCTGTATATACACCTGTGGAGCAATTTTAAACGATCTTTCA	4839
Oy	1213	GLYVALILEUSERPROASNTYRPROGLUBPROTYRPROPROGLIYVSGLIYASPTTP	1232
Db	4840	GGCTTTATCTTTCCACAAACCTCCCTCATCAATCCGCATACAGACAGACTGTGACTGG	4899
Oy	1233	LYSEVALTHRVALSERPROASPTYRVALILEALAEUVALPHEAENILLEPHASLEUGLU	1252
Db	4900	ACTATCAACCGTCANAGACACTGTATATCTCTTGGCGTTCATCAGTTTAAAGATGGAA	4959
Oy	1253	PROGLIYTRAPRAPHLEUHEILETYRASPGLYTRAPASERLEUSERPROLEUILEGLY	1272
Db	4960	CCAAACATATGACTTCTCTATATCTATATATGACACAGCAATTAATAGCCCACTGATGG	5019
Oy	1273	SERPHETRYGLYSERGLINLEUPROGLIYARGILEGLUSERSERASNSERLEUPHELEU	1292
Db	5020	AGTTTTCANAGACAGAACTTACAGAGAGAAATAGAAACACACTCAAAATCAATCATTTGG	5079
Oy	1293	ALAPHEAUSERAPRALSERVALSERASNAAGLYPHEVALILEASPTYRTHGLUBSN	1312
Db	5080	GCTTTTCGAGATGATGATCTGTATAGTTACCTGATTTTCATCGAATACAAAGCAAA	5139
Oy	1313	PROARGUSERCYRPHASPROGLYSERILEYASNGLYTRAPVALIGLYSERASP	1332
Db	5140	CTGGAGAGTCTCTGTTGATTCACAGCAATTAATGATGGCACACAGACTTGGAAATGGAT	5199
Oy	1333	LEUPLYLEUGLYSERSERVALTHRYTRYCYSHASGLIYTRYGVALIUGLYTHR	1352
Db	5200	TATAAATTAAGGATCAACAGTCACTATTAAGTGTGATGCTGGTTATGTTCTTCAAGGTTAT	5259
Oy	1353	SERTHLEUSERCYELILEUGLYPROASGLIYUSPROVALTRPASPAPROARGPRO	1372
Db	5260	TCAACACTCACCTGTATATCATGAGATATGAAAGACCTGGATGGAAATAGACCTTGCCA	5319
Oy	1373	VALCYETHRALAPROCSGLIYGLINTYRVALIGLYSERAPRGLYVALILEUSERPRO	1392
Db	5320	AGTTGTCACTGGCCCTGTGGAAAGTCTTCAACAGGTTTAAAGGACACTGTTCTTACACA	5379
Oy	1393	ASNTYRPROGLINASNTYRTHRSERYGLINILECYSELEUTYRPHLEVALTHVALPROLYS	1412
Db	5380	AACATATCCAAAAAATTAACAGTGGGACATATAATGTGTTATTCATACGACTTCCAAAG	5439
Oy	1413	ASPTYRVALVALPHEGLIYGLNPHLEALAPHEPHEHETHRALAEUASAPRVALIGLU	1432
Db	5440	GAGTTTGAGTGTTGGCCAGATGTTGATATTTTCCAGATACATCCACGATGTTGGTAG	5499
Oy	1433	VALHISAPRGLYHISASERGLNHSERTHRELEUSERLEUSERGLYSERHISETHR	1452
Db	5500	GTTGATGATGGCCCAACTCAGCAATCTTCTGTATATCTTCCCTCCAGATCCCATTTCA	5559
Oy	1453	GLYGLUSERLEUPROLEUALATHRSERASNGINVALLEUILELYSPHESERIALYVSGLY	1472
Db	5560	GGAGATACATTCACATGATTCAGATTAATCAGATCAACAATTCGATTTACTTCACTGGGA	5619
Oy	1473	LEUALAPROALAPRGLYPHEHISAPHEVALTRYGVALALVALPROARGTHRSERALATHR	1492
Db	5620	CCAAATTAACAGCTAAGGGAATTCATCTTGTTTACCAAGCTGTTTCCAGAAACAAGTTCTP	5679
Oy	1493	GLINCYSERSERVALPROGLUBPROARGTRYGLIYUSARGLEUGLYSERASPSPHSEVAL	1512
Db	5660	CAATGCAATCTGTGCTCTGAACCAAGATTCGGAAAGAAATAATGGCAATGGAATTTGCACTC	5739
Oy	1513	GLYALALILEVALARGPHEGLIYCYASNSERGLYTRYALAEUGINGLYSERPROGLIULE	1532
Db	5740	GGTTCAATCGGTTCTTTGATGATTTGATTCAGAGATATATCTTCATGATTCATAGCAATT	5799
Oy	1533	GLUCYSEUPROVALPROGLYALAEUVALAGINTRPASENVALSERIALAPROTHCYEVAL	1552

Db	5800	AGGTGTGAACAGTGGCCCAATTCCTTTGGCCACAGTGAATGATTCTTCACTTATGTAATT	5859
QY	1553	ValProCysGlyGlyAsnLeuThrGluArgGlyThrIleLeuSerProGlyPhePro	1572
Db	5860	GTGCCCTGTGGTGAATTTTAACTAACAAGCGAAAGGAGCTATTTGTGACACCTGGATACCTT	5919
QY	1573	GluProTyrLeuAsnSerLeuAsnCysValTriblyrIleValValProGluGlyValGly	1592
Db	5920	GAGCCTTAATGACAAACAACTGGAATGTGTGTGGAAAGTACACAGTCCAGAGAGGAGCTGGC	5979
QY	1593	IleGlnIleGlnValValSerPheValThrGluIlnAsnTyrAspSerLeuGluValPhe	1612
Db	5980	ATTCAAGTGCAGATGTTGATGCTTTCCTTCAACAGACATATTTGGAGTTCTCTGGACTTTAT	6039
QY	1613	AspGlyIleAspAsnThrValThrMetLeuGlySerPheSerGlyThrThrValProAla	1632
Db	6040	GATGGGGGAGACAAATGCTCCAAAGACTTGGAGATTCATTCAGAACAAACATACCCCAT	6099
QY	1633	LeuLeuAsnSerThrSerAsnGlnLeuTyrLeuHisPheTyrSerAspIleSerValSer	1652
Db	6100	CTTTTGAATGATGCTTAATCTGATCTAATCTTAATTTTCAATGACATCAGAGTTCT	6159
QY	1653	AlaIleGlyPheHisIleGluTyrLeuThrValGlyLeuSerSerCysProGluProAla	1672
Db	6160	GCTCAGAGATTCATCTTGAATACACAGCAATGTTGGATTCTGTCTCTGTAACCAACA	6219
QY	1673	ValProSerAsnGlyValIleValSerThrGlyGlyValArgTyrLeuValAsnAspValValSerPhe	1692
Db	6220	ACTCTTGACAGTGAATTAATAATTGGAGACAGATATATGTTGGTAGATGTATCTCTT	6279
QY	1693	GlnCysGluProGlyTyrAlaLeuGlnGlyHisIleHisIleSerCysMetProGlyThr	1712
Db	6280	CAGTGTATCAGAGATATCTCTTCACAGGTCACCTCACAATTACATGATATGCCAGAGCT	6339
QY	1713	ValArgArgTyrPheAsnTyrProProCysIleIleGlnCysGlyGlyThrValGlu	1732
Db	6340	GTAAGAAGATGAATATTCACATCCCAATTTGTTAGTCACTGAGTGGTGTATGTCACA	6399
QY	1733	GluMetGlnGlyValIleLeuSerProGlyPheProGlyAsnTyrProSerAsnMetAsp	1752
Db	6400	GACTTCAGTGTGTGATCTCAGTCCAGTCCAGGTTCTCTGGAACATATCCACAGAGTTTGAAT	6459
QY	1753	CysSerTriblyrIleAlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPhe	1772
Db	6460	TGCACATGACAAATAAATCTCCACCATGAGTTTGGTGTACATCTCCAGTTTGTAAATTTT	6519
QY	1773	SerThrGluProAsnHisAspTyrIleGluIleArgAsnGlyProTyrGluThrSerArg	1792
Db	6520	TCTACAGAAACCATATCATGATTATTTGGAGTACGAAGATGATCTCAGAAACATGATCT	6579
QY	1793	MetMetGlyArgPheSerGlySerGlnLeuProSerSerLeuLeuSerThrSerHisGlu	1812
Db	6580	GTTATTTGGCCGGCTTAATGATGTCCTCAAAATACATCTTCTTATTCACACACCACTGAA	6639
QY	1813	ThrThrValTyrPheHisAspAspHisSerGlnAsnArgProGlyPheValLeuGluTyr	1832
Db	6640	ACCAAGCTATATTTTCAACAGTACATATTCACAAACAAACAAAGGTTTCAATATGTATAC	6699
QY	1833	GlnAlaTyrGluLeuGlnGlnCysProAspProGluProPheAlaAsnGlyIleValArg	1852
Db	6700	CAACCCATATCAGTTGCAAAAGCTGTCATGATCCAGCCCGTTTCGAAATGTGTTGTATTT	6759
QY	1853	GlyValIleGlyTyrAsnValGlyGlnSerValThrPheGlnCysLeuProGlyTyrGlnLeu	1872
Db	6760	GGTAAATGATTTTACTGTGGGTCAAAACCATTTCAATTGAAATGTTTCCACAGATACACATTA	6819
QY	1873	ThrGlyHisProValLeuThrCysGlnHisGlyThrAsnArgAsnTyrAspHisProLeu	1892
Db	6820	ATTGGAATTCAGCTTCACATGCTTCACAGGATCAGTGTATTTGGAATCATCTCACTT	6879
QY	1893	ProIleCysGluValProCysGlyGlyAsnIleThrSerSerAsnGlyThrValTyrSer	1912

```

Db      6880 CCAGAGTGTGAGCTCTTTGGTGGGAATATACTGCATGATGACCATTTATCT 6939
Qy      1913 ProGlyPheProSerProTyrSerSerSerGlnAspCysValTyrPheIleThrValPro 1932
Db      6940 CCTGGGTAATCTCGATGAATATCCAAATCTTCAAGATGTTGTTGGCTTGTGAAGATACC 6999
Qy      1933 IlegIhAsgIyValArgLeuAsnLeuSerLeuLeuGlnThrGluProSerGlyAspPhe 1952
Db      7000 CCGGGAATGGCATCTACATCAATTTTACTGTCTTCAACAGAACCAATATATGATTTTC 7059
Qy      1953 ILeThrIleTyrPaspGlyProGlnGlnThrAlaProArgLeuGlyValPheThrArgSer 1972
Db      7060 ATTACTGATGGATGGACAGACCAAAATTCACCTCAGATCGGTCTGATCAGTGCGCAAT 7119
Qy      1973 MetAlaIySerThrValGlnSerSerSerAsnGlnValLeuLeuLysPheHisArgAsp 1992
Db      7120 ACCGTTTGGAAATCAGTCTACAGTATCTTCAATCAAGATCTTCAATCAAAATTCACAGTAT 7179
Qy      1993 AlaAlaThrGlyGlyIlePheAlaIleAlaPheSerAlaTyrProLeuThrLysCysPro 2012
Db      7180 TTCAACACAGTGGCTTTTGTGCTGCTCATCATCAGCCTTATCAACCTAAGGGTGTGCCAA 7239
Qy      2013 ProProThrIleLeuProAsnAlaGluValValThrGluAsnGluPheAsnIleGly 2032
Db      7240 CCTCCACCACTGTGCCCAATGCTGAATTTTGAAGGAGATGATGAATTTGAATTAAGGT 7299
Qy      2033 AspIleValArgTyrArgCysLeuProGlyPheThrLeuValGlyAsnGluIleLeuThr 2052
Db      7300 GATATTATTAGTATGATGATGCTCTTCCAGATTTTACTTGAATGCAATTCGATCGACG 7359
Qy      2053 CysAlaIleuGlyThrTyrLeuGlnPheGluGlyProProIleCysGluValHisCys 2072
Db      7360 TGCAGATTAGAGAGACACGTCGACATGATGATGAGACACCTCAGTTTGTCAAGTGTCTGT 7419
Qy      2073 ProThrAsnGluLeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTyrProGly 2092
Db      7420 CTCGCCAATGATTAATCGCGTATGATTTCTACGAGTCAATATGAGCCCTGATATCCCTGAC 7479
Qy      2093 SerTyrProGlnPheGlnThrCysSerTyrPheValArgValGluProAspTyrAsnIle 2112
Db      7480 AGTTAACCAATCTTCAATGATGTGATGAGACATTCACCTGAGAAAGGTTATTAATATC 7539
Qy      2113 SerLeuThrValGluTyrPheLeuSerGluLysGlnTyrAspGluPheGluIlePheAsp 2132
Db      7540 ACCATGTTGTAGATTTCTTCACAGACGAAAGAAATTTGATGTTCTTCAGGTGATGAT 7599
Qy      2133 GlyProSerGlyLysSerProLeuLeuLysAlaLeuSerGlyAsnTyrSerAlaProLeu 2152
Db      7600 GGACCAATATATCAAAATCCAGTCTTATTTCCCTCAGTGGGATTAATCATCTGCTTTT 7659
Qy      2153 IleValThrSerSerSerAsnSerValTyrLeuArgTyrPheSerSerAspHisAlaTyrAsn 2172
Db      7660 AATATTAACAAGCAATGCTGATGAAGTATTTCTTCAGTGTGACAGATCATGTGCATATAC 7719
Qy      2173 ArgIyGlyPheLysIleArgTyrSerAlaProTyrCysSerLeuProArgAlaProLeu 2192
Db      7720 AAAAAAGGCTTCGGGAAATATATAGCTTTCTACTGTAGTACACCAAAATCCCAACCT 7779
Qy      2193 HisGlyPheIleuGlyGlnThrSerThrGlnProGlyGlySerIleHisPheGlyCys 2212
Db      7780 CATGATATATATCATGCTCAGACAGAGTGGGACCTTAACAGTGTGCTGTTGGGCTGT 7839
Qy      2213 AsnAlaGlyTyrArgLeuValGlyHisSerMetAlaIleCysThrArgHisProGlnGly 2232
Db      7840 GATCGAGGATTCGACCTTGTGGAAAAAGCAGTGTGTGCGAAGAACTTCCATATGGG 7899
Qy      2233 TyrHisLeuThrPheSerGluAlaIleProLeuCysGlnAlaLeuSerCysGlyLeuProGlu 2252
Db      7900 TATCATGATGATGATGCGCAGTCCCTGCTGCAAGCAATTTCCGTGGGATTCCTTAA 7959
Qy      2253 AlaProLysAsnGlyMetValPheGlyLysGluTyrThrValGlyThrLysAlaValTyr 2272
Db      7960 GCTCCAAACAATGAGAGAAATATCAACAACAGACTATTTGTAGGAACGGAGATTACCTAT 8019

```

```

Qy      2273 SerCysSerGluGlyTyrHisLeuGlnAlaGlyAlaIleThrAlaGluCysLeuAsp 2292
Db      8020 TTTTGAATGATGAGATATGATGATGATCAAAAGAACTCACTACATGCTGATAGCCAATCA 8079
Qy      2293 ThrGlyLeuThrPheSerAsnArgAsnValProProGlnCysValProValThrCysProAsp 2312
Db      8080 GATGGAACATGGAGCATATATACAAGACCCCTCGCTGTGTGTGTATCAATGCTCCAAAC 8139
Qy      2313 ValSerSerIleSerValGluHisIleArgTyrArgLeuIlePheGluThrGlnTyrGln 2332
Db      8140 ATCAATCTCTTATCTTGTGAACATGAGATGGCGAATTGTGAAGGCTCCCATATATGA 8199
Qy      2333 PheGlnAlaGlnLeuMetLeuIleCysAspProGlyTyrTyrTyrThrGlyGlnArgVal 2352
Db      8200 TACAAACCAAGATGATGTTTACGTGTGACCTGTGTATCATGAGACTAGTCTCCTCC 8259
Qy      2353 IleArgCysGlnAlaAsnGlyLysThrPheSerLeuGlyAspSerThrProThrCysArgIle 2372
Db      8260 ATCGAATGCTCTCTAATGATGATCTTGAGTGGAGAAATGAAAGCCATATTTGCCAAAT 8319
Qy      2373 ILeSerCysGlyGluLeuProIleProProAsnGlyHisArgIleGlyThrLeuSerVal 2392
Db      8320 ATTCTCTGTGAGAACTACCTACACCTCCAAATGGAATATGATGGAATCAAACTTCA 8379
Qy      2393 TyrGlyAlaThrAlaIlePheSerCysAsnSerGlyTyrThrLeuValGlySerArgVal 2412
Db      8380 TATGGCTCAACAGCTATCTTATCTTACCTGCACTGGGATTCATGCTTGTGGCTCTGCTTGA 8439
Qy      2413 ArgGluCysMetAlaAsnGlyLeuThrPheSerGlySerGluValArgCysLeuAlaGlyHis 2432
Db      8440 ACGAATGCTTTCTCTCAGTCTTTTGAGTGAATGTGAACCAAGATGCTTACGGGTAT 8499
Qy      2433 CysGlyThrProGluProIleValAsnGlyHisIleAsnGlyGluAsnTyrSerTyrArg 2452
Db      8500 TGTGAAATTCAGAACTGATGTGATGTGATGATGATGAGAAATATATGATATATAGA 8559
Qy      2453 GlySerValValTyrGlnCysAsnAlaGlyPheArgLeuIleGlyMetSerValArgIle 2472
Db      8560 GACACAGTGTATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8619
Qy      2473 CysGlnGlnAspHisHisIleTyrSerGlyLysThrProPheCysValProIleThrCysGly 2492
Db      8620 TGTCAACAGATTCACAAATGCTGTGCTGCTCCATCTGTGTGCTGCTGCTGCTGCTGCT 8679
Qy      2493 HisProGlyAsnProValAsnGlyLeuThrGlnGlnGlnPheAsnLeuAsnAspVal 2512
Db      8680 CACCTGTGATGCTCAATTTATGAGAAACAAAGTGAATGGGTCAACTTAAATGATGTG 8739
Qy      2513 ValLysPheValCysAsnProGlyTyrMetAlaGluGlyAlaAlaArgSerGlnCysLeu 2532
Db      8740 GTAACATTTCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8799
Qy      2533 AlaSerGlyGlnThrPheSerAspMetLeuProThrCysArgIleIleAsnCysThrAspPro 2552
Db      8800 GCCAACAGACAGTGGAGCATCTCCACATATGCAAAAGGTGCAACTGTGTGATCTCT 8859
Qy      2553 GlyHisGlnGluAsnSerValArgGln-----ValHisAlaSerGlyProHisArg 2569
Db      8860 GGAATTCAGCCAAATCTCAAAAGAAAGTAAATAGAACATGGAAT----- 8907
Qy      2570 PheSerPheGlyThrThrValSerTyrArgCysAsnHisGlyPheTyrLeuLeuGlyThr 2589
Db      8908 TTTACTTAAGGCACTGTGATATTTCTAGCTGCAATCTGATATTTTATTTTATTTGATCT 8967
Qy      2590 ProValLeuSerCysGlnGlyAspGlyThrTyrAspArgProArgProGlnCysLeuLeu 2609
Db      8968 TCAGTTTGTATGTCAACCAAAATGAGCAATGGGCAAACTTTTACAGATGATATCATG 9027
Qy      2610 ValSerCysGlyHisProGlySerProProHisSerGlnMetSerGlyAspSerTyrThr 2629
Db      9028 ATTACTGTGACACCTCGGCGTTCTCTTAATGACAGTCTGTGTGCGAAGATATATCT 9087

```

Qy	2630	ValGlyValAlaValAlaGlyYSerCysGlyIleGlyYValArgThrLeuValGlyYAsnSerThr	2649
Db	9088	TTTGGGCTACTGTTCACTTCTCTGACACAGAAAGGCTTCCCTTTTAGGCACGTATCA	9147
Qy	2650	ArgMetCysGlyLeuAspGlyYHisETripThcGlySerLeuProHisCysSerGlyYThrSer	2669
Db	9148	AGAACCTGCCAATTGAATGGCACTTGGAGTGAGTACAACTCTATTCTTCAGGTGATGCT	9207
Qy	2670	ValGlyValCysGlyYAspProGlyYLeProAlaHisGlyYLeArgLeuGlyYAspSerPhe	2689
Db	9208	ACTGGGACATGTGGCGCATCCACAGTACTCCCGCCATGCTGCTTAGACACGAAAGCAATTTC	9267
Qy	2690	AspProGlyYThrValMetArgPheSerCysGlyLysIleGlyHisValLeuArgGlyYSerSer	2709
Db	9268	AGAACTAAAGTACTGTACGTTAGCTTGTAICTGGTATCATCTTCATCTTCATGGCTCGAA	9327
Qy	2710	GluArgThrCysGlnAlaAsnGlySerTripSerGlySerGlnProGluCysGlyValIle	2729
Db	9328	GAAAGAACATGTTTAGCTAATGCGAGTTGGACCGGAAAGCGACCGCAGTGTCCAAAGCTGTG	9387
Qy	2730	SerCysGlyYAsnProGlyYThrProSerAsnAlaArgValAlaPheSerAspGlyYLeuVal	2749
Db	9388	CAGTGTGGTAACTCCAGAACACACAGCAATGGAAAGTCTTCGCAATTGATGGACAAACA	9447
Qy	2750	PheSerSerSerIleValYTrGluCysArgGluGlyYTrYTrAlaThrGlyYLeuLeuSer	2769
Db	9448	TTTTCTAGTTCAGTCATTTATTTCTGTGATGGAGGAAATACATCTTCTTGACCTTCAGTT	9507
Qy	2770	ArgHisCysSerValAsnGlyYThrTripThcGlySerAspProGluCysLeuValIleAsn	2789
Db	9508	AGACAGTGTACACAGCCAAATGAAACATGTGCTGGAACTTACTTACTGAATCAATATACGT	9567
Qy	2790	CysGlyYAspProGlyYIleProAlaAsnGlyYLeuArgLeuGlyYAsnAspPheArgYTrAsn	2809
Db	9568	TGTGGAGACCCAGAGTATCCAGCCAAATGACATGATGAGATGATGATTAATGTGGTTGGA	9627
Qy	2810	LysThrValThrYTrGlnCysValAlProGlyYTrMetMetGlySerHisAlaGlyValSerVal	2829
Db	9628	CAAAATGTTTCTTACATGTGTCCACCCAGGCTACACGATGGAAATGGATGGCTCCAGATTC	9687
Qy	2830	LeuSerCysThrLysAspArgThrTripAsnGlyYThrLysProValCysLysAlaLeuMet	2849
Db	9688	AGGACTGTACAAATTAATGACACATCGAGTGGAGTAATGCCAACTGTAAAGCTGTTACC	9747
Qy	2850	CysAllybProProLeuLysIleProAlaAsnGlyYLeuValAlaGlySerAspPheMetTripGly	2869
Db	9748	TGCCCACTCTCCCCACAGATCTTAATGAAGAACGCTGGAAGGAACAATTTTCAGCTGGGCG	9807
Qy	2870	SerSerValThrYTrAlaCysLeuGluGlyYTrGlnLeuSerLeuProAlaValPheThr	2889
Db	9808	TTTATGTATTACCTACATCTGTTCTCCAGAGCTATAGCTATCTTCCCTGCTGTTTGGACC	9867
Qy	2890	CysGluGlyYAsnGlyYSerTripThcGlyGluLeuProGlnCysPheProValPheCysGlyY	2909
Db	9868	TGTGTAGGAATGGTACTGGAGTGGTGAATGACCGCAGTGCCTTACCAAACTTTGTGGT	9927
Qy	2910	AspProGlyYAlaProSerArgGlyYArgArgGluAspArgGlyYPheSerYTrArgSerSer	2929
Db	9928	GACCTGGTAACTCTGCCCAAGAAAAGAGAGAGCAAAAGCTTTATATACCACTACAG	9987
Qy	2930	ValSerPheSerCysHisbProProLeuValLeuValGlySerProArgArgPheCysGln	2949
Db	9988	GTTTCATTTACGCTCAATTTTCTTTCAATATTAGTGGAGTACAGACCAAAATATGTCAA	1004
Qy	2950	SerAspGlyYThrTripSerGlyYThcGlnProSerCysIleAspProThrLeuThrTrnCys	2969
Db	10048	GCAAGTGGCACTTGGAGTGGTTCATCACTTACATGCAATAGGCGCTTACCAAACTCTTGT	10107
Qy	2970	AlaAspProGlyYAlaProGlnPheGlyYIleGlnAsnAsnSerGlnGlyYTrGlnValGlyY	2989
Db	10108	GAAGAACCCAGGTGTGCTCCGCGACAGATCTCAGAAACAATACATTCCGATTTTCAAGTAGCA	1016
Qy	2990	SerThrValLeuPheArgCysGlnLysGlyYTrLeuLeuGlnGlySerThrThrArgThr	3009

Dd		10168	AGTCTTTACAGTTCCATTCGCAAAAAGACACACTTCTCCAAAGGTCTACAAACCACC	1022
Oy		3010	CysleuProAsnLeuThrTrpSerGlyThrProProAspCysValProHisHisCysArg	3029
Dd		10228	TGCCTCCCTCATCTACGTGAGTGGAATTCACCGCTGAATGATACCCACAGCTGTAA	10287
Oy		3030	GlnProGlnThrProThrHisAlaAenValGlyAlaLeuAspLeuProSerMetGlyTrp	3049
Dd		10288	CAGCCACAACACTCCTGCTCATGCAAAATGTCGTAGGAGATGACCTTCATCTCATGGGTAT	10347
Oy		3050	Threuleile 3052	
Dd		10348	ACACTGATT 10356	
RESULT 10				
LOCUS	ABI14604	13148 bp	mRNA	linear PRI 29-AUG-2003
DEFINITION	Homo sapiens mRNA for CSMD3 protein isoform 1, transcript variant			
ACCESSION	ABI14604			
VERSION	ABI14604.1	GI:34330130		
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1 Shimizu A., Asakawa S., Sakaki T., Yamazaki S., Yamagata H., Kudo H., Minoshima S., Kondo I. and Shimizu N. A novel giant gene CSMD3 encoding a protein with CUB and sushi multiple domains: a candidate gene for benign adult familial myoclonic epilepsy on human chromosome 8q23.3-q24.1 Biochem. Biophys. Res. Commun. 309 (1), 143-154 (2003)			
JOURNAL	2 (bases 1 to 13148) Shimizu N., Asakawa S. and Shimizu A. Direct Submission Submitted (11-JUL-2003) Nobuyoshi Shimizu, Keio University School of Medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan (E-mail:nshimizu@med.keio.ac.jp, Tel:81-3-3351-2370, Fax:81-3-3351-2370)			
MEDLINE	22824680			
PUBMED	12943675			
FEATURES				
source	1. 13148 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="8" /map="8q23.3" /tissue_type="brain" /dev_stage="fetus" 1. 13148 /gene="CSMD3" 160..11283 /gene="CSMD3" /product="CSMD3 protein isoform 1" /protein_id="BAc82443.1" /db_xref="GI:34330131" /translation="MKGIRKGESRAKSPKEPGRCKAKCGRLDFTLMKKMGISGF TFNNVFLVLIVSGCVKGFITYCGGTGLNGLTIESPGFPGYRGNACTWIIIAEBNR IQIWPFSALIEBYDYLSITDGHPHTNRTLGTFLPPPTSTKSVSLSLTDF VSAHGFRVYYEELOSSSGKNPVGPVGVIYGRFDVGDRIKRYSCVTGYILDHPLDT IASNVNTASMDPVPICRAPNDGCMRGSSGCISSPSFPENYHNADQTMIVAEPC DTSLIFTDQMEKYDIYEISEEPPTIMLGAMIPPIIISKMKLRHPTDSNR YRFSAAYOGSSTLTHRTSTGELEBNRRITTAIAASTPAADVTSVAIVTRHISE EQGVQVTSJRNGLDPNTSDKGLSPHADTOSTRRPRABQIERTKIALAVTHRKYK AIDPFGRGFLTPGKDNKNKFSLINSGKITASNLCPPDGEENGKRIGSDLSSTV QFCDDVDYLOGAKSITCORIAEVFAAMDHRPVCVKTCGNSLOGPSSTFPSNPPE QYSNMCAWTVITAVNTAVVICINFEPEPLETGYDTLTGDDGEGVDPPTVALVLTGS RPVDILVASMSQMHLQTDVESGVSQFVRANYKEIKESKCDPPTLYVILIRGDDFSN RDVLRBCQFGFELIGEKSIIVOCENNOMANIPIICPLSNFTAMPGLVLSPDPEG			

YANNLCIMTIIISDPGRSHLSPNDPDLSEQDFELAVKDGSPESPIIGFTGAESVS
 HLTNSHILRLFEQADHSMGSRGNITNTGNHECPDGPPIARBPFGNFOGSSI
 SVTCEBEPKTKQETETTCILMDGKMMWSGLIPKCGAPCGHGRAPBGLISBPMXP
 YKDISNCWVIEARBSHRIKTPRPEROTELAVDYLVDHNGENLSPILGSTNGVQV
 FLBSSSNRYILFTTDSNRNNRFRKHIESVTAVYCLDDEGLVHVRGRGHPEIS
 TVPSCDGYSRLSHEBLLCEKNHWSHPLPTCDALCGDGRGSGTLLBPGVEFP
 NSLNCWTAVDVTHKGQVFNPHFPHLEHDHLYLITNGSGTFOLALITGSDLPPTN
 AGLYGNFAQLRFISDFSIYEGFNITFSEENLPCEDPDGPQOYSGISGNFGIDL
 TFGSCSGYRLGETSEIICLGGGRVMSAPLRCVABGASATNNNGILLSPNPLANE
 NNHCISYIOVAKGINISARTPHLAGOVLKTYDKDKDTHHLGAFASNRGLTL
 SSTNOLMLENSDTEGDEFGOLVYSFELSHCEDGEPQEPQKISDQHGASTII
 YCNPFTLHSSSLCKWTKERRAMDPLPSCIAECGARGESGRLLSBPYFPD
 NNLKMMMI EVDPNIVSLQFLADTASHDILRWMPGPNEMDLLEIGISDLPBGI
 HSTINIVTIQFDTPFYSIGSGFAIQFSSVATACRDGPNGRNDGPGDPTVAF
 QCDPGELOGEERITCLOENRYFMOPSPVCIAPCGNLTGSSGFLISNPHPHPH
 SRDDMTITVADYVIALAFISFISIPNYPCLTYDPDPSNPLISGPOSKPERE
 SSSMTMLAPRSDSVSYTGPHLYAKALBESCDPBNINMGRBLGNDYLSVTVY
 CDAYVIGSTILTCIMGDDGRPEGNRALPSCNAPCGSRGTSGEVLISNPKNSV
 GHNCVSIYAVEKEFVFGQVFPQTSLHDVEVYDGPQOOSLSLSLGSSESLPL
 SSGNQITIRFSTVGPITAKGFHVEYQAVPRTSSTQCSVEPREGRIGNEFAVGSV
 LFDNPGYILHSGSIAICETVPSLAQMNDSLPCTIAPCGIILTKRGITLSGYEP
 YDNNLCNWKITVEBAGIOVOVSPATENHMSLDPYDGDNNAPRLGSSGTTIPH
 LNSTNNIYANPQSDISVSAAGPHLYTITAGDSCPEPOTPSGIRIGRYVAGV
 SFQDCQYSILOGSHITCWPBPVRNMYPIPTICLAGAMSPSGVILSPGPNYV
 SLDCTWTINLPISFVHALQFVNSTETIHDYLEVRGSESSTVIGRLSGPOIPSL
 ESTHETSLYFSDYSONKQGFHLYOAYLOSCPDPRPRNPFVIGNDPTVQOTISF
 ECPGYTLIGNSALTCIHGVSNNHMLPRLPCALCGNITANNGTISPGYDEYFP
 QDCFMLVRPENGIIYINFTVLOTETPYDTRCOPPDNDSPOTIGSGNTALIESYV
 TSNQILIKHPSDFTSGPFVLSAYOLRTPCOPPPRPNALITTEDBFRICITTYO
 CLBPTLVGNAILTCRGERLQMDGAPVQVCPALBLLDSTGYILSGYDSTYN
 LQNCAMSTVSKGNTIMTFVEFQTEKEFVLVYDGPNIQSPLVLSLSDYSANIL
 TSNQHEVLSVSKADHNNKGFRIYAFYCSPTSPSPHYIISQTCQOLNSVYRAC
 DRGRLVGSASVAVCRKSSYGHAMDAVPACQALISCIPIAPFNGIILHDMRYGRV
 TYFENDGYRLSSKELTTVACOSDGTNNHKTBRCVVTCPSINSFLLHGRMYING
 SHYEYTKVYVSCDPGHHGHPASIECLPMTGSMNERPYCOIISGELPTPBNK
 IGTQTSYGAITFTCDGPMVNSAVRECLSSGMSBSETRCLAGHGBELTVNQV
 IGENYGRDVTVOQNPFRLLISSVARIQQDHNWSGLTSCVAGCGHSPITYRT
 SNGENFNDVTVFSNIGLYMQPTKAOCCANROMWHPMPCVAVNCBIPANRKR
 ESKIEHNFYTVGVPCNPGYFLFGSSVILCQPNQOMKFLPEECIMIDCGHVP
 NAVLSGKTYFGSTVHVSCTGRSLGQSERTCOLNGHWSGPHSGSAGATCGPG
 TPGHGRQESNFPRTKSTVYACDGYILHESERTCOLANGSMWRGSKAVOCCGPG
 TTANGKVRIDGTGTSVYSCMEGIIISGVSROCTANGTSGILPCTIISCDP
 GTPANGKRYDGTVVOQNVSYKCPQGTNELNSRIRTCITNGTMSGEVPCAVCP
 TPPOISNGRLEGNPWFESISYICSEBVELSPAVLTCNGTMSGEVPCAVCP
 DPGIPAGKREGKSFYIOSEVSPSCNFPPLIVGSTRICQADTMSGSPHCIEPOT
 SCENPVPGRSQNNTFGFQVSGVFOHCKGHLLOGSTRTCLPLTMSGLOPEICP
 HSCKQEPITPAHNVGMDLPSHGVTILYTQPGPFLAGGHEHVCSNDUMWGXVPC
 EAGSKILVDRPALGPPSKLSVDDVPAQNTYKWSYNTFRKQKPTMTLYTSRVS
 TGRVNTLSNNMELLISYKQSEARMLARILYIKVPHASVAKMEENMAMDEGVS
 AEPDGYVYQGTIQGKDYGQFQLORLGLNMSGSSSHQPHQTNSSVAIALIVPFP
 ALIPAGEFYLKQRTAPKTYGTCVSHENNQOAFENPMYDITNAKSVGKAVPDP
 NNTVCTMTV"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 13148
 Score: 10591.50 Matches: 1816
 Percent Similarity: 76.65% Conservative: 501
 Best Local Similarity: 60.07% Mismatches: 665
 Query Match: 62.36% Indels: 41
 DB: 9 Gaps: 6

US-10-016-248-2 (1-3104) x AB114604 (1-13148)

QY 36 VallyleuGlnIleuLeuLysSerArgGlyVallyleuLeuMetProSerLysAspAsn 55
 DB 1501 GTGAAAAAGCCCTAGATTAAATCTAGAGATTAAATTTGTTTCCAGGAGAAAGCAAC 1560
 QY 56 SerGlnLysThrSerValleuThrGlnValGlyValSerGlnGlyHisAsnMetCysPro 75
 DB 1561 AGCAACAAAGTTTCTATCTTAATAGAGGAGGATATTAACAGCTTCCAAATTATGCCCCA 1620
 QY 76 AspProGlyIleProGlnArgGlyLysArgLeuGlySerAspPheArgLeuGlySerSer 95

DB 1621 GATCCGAGAACCAACAAAATGGAAGAGAAATCGATTCGATTTTACCTTGATTAACCT 1680
 QY 96 ValGlnPheThrCysAsnGlnGlyLysArgLeuGlnGlySerLysArgIleThrCysMet 115
 DB 1681 GTGCAGCTTCCTTGATGATGAAATTAATGCTTACAGGGCGCAAGAGCATACCTGTCA 1740
 QY 116 LysValSerAspMetThrAlaAlaTrpSerArgPheAlaArgProValCysArgAlaArgMet 135
 DB 1741 CGGATGCTGAAGTTTGTCTGCTTGAAGTATCACAGGCGGTGTGTGAATGTGAAGAACG 1800
 QY 136 CysAspAlaHisIleuValArgGlyLysProSerGlyIleIleThrSerProAsnPheProIleGln 155
 DB 1801 TGTGGGTCTTAATCTTCAAGAGCAAGAGTGTACCTTACATCTTCCCACTTTCCTTCCAG 1860
 QY 156 TyrAspAsnAlaHisCysValTrpIleIleThrAlaIleuAsnProSerLysValIle 175
 DB 1861 TATGACAGCAATGCAACAAATGTCTGTGGTCATCACAGCAGTGAATCAAAATAGGTTATC 1920
 QY 176 LysIleuAlaPheGlnGluPheAsnLeuGlnArgGlyLysArgThrLeuThrValGlyAsp 195
 DB 1921 CAGATTAATTTGAAGAAATTTGATCTGAGATGGCTATGATACCTTGACATTTGGCCGAT 1980
 QY 196 GtlyGlyGlnAspGlyAspGlnLysThrValleuLysThrValleuLysSerGlnAsnAlaCysSerAsp 215
 DB 1981 GGGGGGGAAGTTGAGATCTTACAGACATGCTCCCAAGTG----- 2019
 QY 216 SerProHisThrProGlySerArgIleProGlu-----SerMetSerGlyAspIle 232
 DB 2020 -----CTGACTGGAAGCTTTGTACAGACTGATAGTACGATGAGCATGAGCCAAATG 2070
 QY 233 TrpArgGlnLysTrpThrValleuGlnIleCysArgAspIleSerSerAspAlaArg 252
 DB 2071 TGGCTGCACCTTCAACG-----GACCAAGT 2097
 QY 253 SerGlySerValArgLysSerProLysThrSerAsnAlaValGluLeuValAlaProGly 272
 DB 2098 GTTGATCTGTTGGTTCAAGGTTAACATCAAA----- 2130
 QY 273 ThrGlnIleGlnGlnGlySerCysGlyAspProGlyIleProAlaArgGlyArgGlu 292
 DB 2131 ---GAAATTTGAGAAAGAAAGTTGTGTGATCTGTTACACCTTATATGGAATTAAGAA 2187
 QY 293 GlySerAspPheHisGlyAspThrIleuLysPheGlnCysGlnProAlaPheGlnLeu 312
 DB 2188 GCGCATGATTTTCTAATCGGATGTTTAAAGGTTTAATCCAGTTTGGGTTGAATTA 2247
 QY 313 ValGlyLysValAlaIleThrCysGlnLysAsnAsnGlnTrpSerAlaLysLysProGly 332
 DB 2248 ATTGGAGAGAAATCATGTTGTTCAAGAGAAATACCAATGCTTGCAGAAATATGCCATC 2307
 QY 333 CysValPheSerCysPhePheAsnPheThrSerProSerGlyValValleuSerProAsn 352
 DB 2308 TGTATCTTCCCTGCTGCTCTTAACCTTACCTGACCAATGGAGAACGTTCTTCTCTGAT 2367
 QY 353 TyrProGluAspTyrGlyAsnHisIleuHisCysValTrpLeuIleLeuAlaArgProGlu 372
 DB 2368 TACCCAGAGAGGTATGAAAAATTTAAATTCGATCTGAGAGATAACTCTGATCCAGGG 2427
 QY 373 SerArgIleHisIleuAlaPheAsnAspIleAspValGlnProGlnPheAspPheLeuVal 392
 DB 2428 AGCCGATACATCTTCTTCAATGATGATCTGAGATCCGAGATCCAGTTGATTTCTTGCT 2487
 QY 393 IleLysAspGlyAlaThrAlaGluAlaProValleuGlnLysPheSerGlyAsnGlnLeu 412
 DB 2488 GTTAAAGTGTGATCTTCCAGAAATCCCAATTTCTTGAACCTTTATCTGGGCGTAGAGTG 2547
 QY 413 ProSerSerIleThrSerSerGlyHisIleValAlaArgLeuGlnPheGlnThrAspHisSer 432
 DB 2548 CTTTCCATCTTACTAGTAATAGCACATCTGGAATTTGGAATTTACGCTGACCATCTCA 2607
 QY 433 ThrGlyLysArgGlyPheAsnIleThrPheThrPheArgHisAsnGluCysProAsp 452

```
Db 2608 ATGCAGACGCGTCTTAACATCATCAACACATTGGCATTAATGAATGCCCTGAT 2667
Qy 453 ProcllyalProvalaenqlylyarqPhegllyasPserleuqInleuqlySerSerlle 472
Db 2668 CTTGGAAATACCAATCAATGACGCGGCTTTGGGACCACTTCAATTAAGAAAGTTCAAT 2727
Qy 473 SerPheleuCyasaPgluglyPheleuqlyThrgInqlySerqInleuqlyleThPheVal 492
Db 2728 TCACTTATTTGTGAAGAAAGATTATTAACCCAGGAGACAAATTAATCAATGATTT 2787
Qy 493 LeuylagluglySerValaJlTPAsnSerlaValleuqlyCyasgluaJlProCyasglly 512
Db 2788 CTTATGATGGAAGAAAGTATGAGTGAATTCGAATCCAAATATGAGAGCCCATGAGT 2847
Qy 513 GlynleuThrsPserProSerqlyThrlleuSerProqlylTPProqlyPheThlyls 532
Db 2848 GGCATTTTTCAGATCCACAGGAGGAGTCTCTCACACAGATGCCAGGATCTACAA 2907
Qy 533 AspaJlaleuSerCyasaJlTPValJllegluJlagnProqlylTPProJlaleylleTh 552
Db 2908 GACTCTTGAATTTGTAGTGGGTATGAGTGAAGCTGAACCTGACACTCTATCAAAATTA 2967
Qy 553 PheAspArgPheylsrhrgluValaAsnTyraPThrleuqlyValaArgspglYArgThr 572
Db 2968 TTGAAAGATTTGAGACTGAACTGAATATGATGTTGTGAAGTTCAATGATGGCCAA 3027
Qy 573 TyrsPrlaProlleuqlyleqlyValaTyrlleqlyThrgInvalProqInPheleuJlSer 592
Db 3028 CTCTGTGACCCCTGTGATCTTACATGACACCCAAAGTCCAGTTTCTATTTAGT 3087
Qy 593 ThrsPasnTyrlleuTyrlleuPheSerThrsPlysserThrsPAspJlleglyPhe 612
Db 3088 AGCGATATTTTATATACCTTCTATTTACACAGACACAGCTCGTTCCAAATATGATTC 3147
Qy 613 GluleuArgTyrgJlurThrlleThrleuqInSerAspHisCyaleuAspProqlylIlePro 632
Db 3148 AAGATTCATTTATGAAGAGTGAAGTGAACAGTATCTGTTGGACCCCTGGATACCT 3207
Qy 633 ValaenqlygluaqghleqlyAsnaAspPheTyrlValglYalaleuValThrPheSerCy 652
Db 3208 GTATATGCGCCCTCCCTATGCTGATGATTTCTCCATGCTGCTACTGTTCTATTTAGTGT 3267
Qy 653 AspserqlyTyThrlleuSerAspJlgluProlleuqlyCyasgluProlleuPheqInTP 672
Db 3268 GATTCAAGATACAGCTTGAAGTGAAGAGCCCTTCTATGCGAAATAAACACCTGGTG 3327
Qy 673 SerArgJlaleuProlleuSerCyasgluJlaleuCyasglYlPheleqInqlySerSerqly 692
Db 3328 AGTATCCACTTCCACACTGCTGATGATTAATGTGAGAGATGTAGAGGCCCTAGTGA 3387
Qy 693 ThrlleuSerProqlyPheProlleuPheTyrlProlleuAsnleuAsnCyethrTPle 712
Db 3388 ACAATCTTATCACCTGCTTACCGGAAATTTATCCAAATCTCGAATTTGATGACATGACT 3447
Qy 713 lIleqInThrsPheThrlleqlylyglYValPhePheThrsPheThrlleqlyleuqlySer 732
Db 3448 GTTATGATACCCAGAAAGAGTGTGAGTTCACATTCACACTTTTCATTTGAAAGAC 3507
Qy 733 GlynleuAspTyrlleuLeuJlThrgluJlAsnqlySerPheThrgInProlleuArgJlInleu 752
Db 3508 CATCATGACTTACTACATGATCAGAGATGGAGTTTACCACACCTGGACGCGCTG 3567
Qy 753 ThrlleuSerArgleuProlleuProlleuSerlaqlyleuTyrglyAsnPheThrlaJlIn 772
Db 3568 ACTGATTCAGATCTTCTCCCAACATCATGCTGTCTATGGAATTTCCAGGCTCA 3627
Qy 773 ValArgPheleuSerAspPheSerMetSerTyrgluqlyPheAsnleuThrPheSerqly 792
Db 3628 TTGCTTTCATTTCAATTTTTCATATATCAATAGATTTAACTAACATTTCTTGA 3687
Qy 793 TyraPleuqInProCyasgluqInProqlyValaProJlATyrsPrlleArglyleu 812
Db 3688 TATAACCTTGAACTTGTGAAGATCTGCGATTCCTCAATATGTGTAGCGAATGGGTT 3747
Qy 813 GluPheglYalGlyAspThrlleuThrsPheSerCyasPheProqlylTPArgleuqly 832
Db 3748 AACTTGGAGTGTGTGACCTTGAACCTTCTCATGCTCTTGGGTTATCACTGGAAG 3807
Qy 833 ThrlaArgleThrCyaleuqlyglYArgArgArgleuThrsPserProleuProArg 852
Db 3808 ACATCAAGATCATCTCTCTTGGTGGTGGCCGAGAGATGTGAGTGCACCTCTGCCAAG 3867
Qy 853 CyasaJlaleuCyasglYasnservalThrglyThrgInqlyThrleuLeuSerProAsn 872
Db 3868 TGTGTGCTGAATGTGTGATCTGCAACAAATATGAAGAAATTTGTCTGCTCCAA 3927
Qy 873 PheProValaAsnTyraAsnaAsnHisglYCyelIleTyrsPrlleqInThrgInProqly 892
Db 3928 TATTCCTCAACTGAATGAACCAACAGAAATGATTAATGATTCAGTTCAAGCAGA 3987
Qy 893 lYasglYllegluInleuysaJlAArgJlAphglYleuSerqlyAspValleuYVal 912
Db 3988 AAGGAAATCAATTTTCAGCCAGAACATTCATTAGCAGAGAGATGTTCTTAAGAT 4047
Qy 913 TyraPglYasnaAsnSerlaArgleuqlyValaPheSerHisSerqlyMet 932
Db 4048 TATGATGAAAGAAATAAACAGACTCATCTAGTCTTTTACTGTGATCATATGCGC 4107
Qy 933 GlynalThrlleuAsnSerThrsPserSerleuThrlleuAspPheleThrsPrlaJl 952
Db 4108 GGAATGACCTTGTATGATCTTCAAAATCACTGCTGATTAATTAATTCGATGAA 4167
Qy 953 AsnThrsPserqlyPhegluleuHisPheSerPhegluleuJlaleuYCyasgluAsp 972
Db 4168 GCGACAGATGAAGCTTCACTTGTGTATACCAAGTTTGAACCTCACACTGGAAGAT 4227
Qy 973 ProqlylThrProlleuPheglYTyrlleValaHisAspJlglYHisPhealaglySerSer 992
Db 4228 CTTGGACATTCACAAATTTGATACAGATACAGACACCAAGCCACTTGTGTAAGACC 4287
Qy 993 ValSerPheSerCyasaPProqlylTPYrsPleuArglySerqInleuqlyleuCyaleu 1012
Db 4288 ATCATTTATGATGAATCCAGCTACCTCTCACAGAAAGTACCTTCTCAAGTGA 4347
Qy 1013 SerqlygluaArgThrTPAspArgProlleuProThrsCyasaJlaleuCyasglY 1032
Db 4348 ACAGGAGAGAAAGGATGAGATACCTCTGCTCTCTGATGTGATGAGAGT 4407
Qy 1033 ThrsValaArglygluValaSerqlyglYValleuSerProqlylTPProJlaleuProqly 1052
Db 4408 CGTTTAAAGAGATCATCAGAGAAATCTTATCTCTGCTATCTTTCCATATGAC 4467
Qy 1053 HisAsnleuAsnCyelIleThrlleqlygluaJlaleuqlyCyethrlleqlyleuHis 1072
Db 4468 AATTAACCTGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4527
Qy 1073 PheleuValaPheAspThrgluqlyValaHisAspValleuArgleThrsPAspJlProval 1092
Db 4528 TTTCTTCTTTTATGATGAGAGATCATCATGATATCTCCGATCTGGAGCGGTCCACA 4587
Qy 1093 GluSerqlyValleuLeuqlygluLeuSerqlyProJlaleuProlleuAspPheSer 1112
Db 4588 GAATATGATATGCTTTTAAAGAAATTAAGATCTCTTATCTCTGAAGAAATTCATAGC 4647
Qy 1113 ThrsPheAsnSerValaValleuqInPheSerThrsPhePheThrsPrlleqlyglYlPhe 1132
Db 4648 ACCCTCAATATGATCAATCACTGATGATGATGATGATGATGATGATGATGATGAT 4707
Qy 1133 AlaJllegluPheSerValSerThrlaThrsCyasaAspProqlylIleProqInAsn 1152
Db 4708 GCAATTCAGTTTCAAGTCTGTGTCACATCGCTGTGTGATCCAGAGGCTCCCATAT 4767
Qy 1153 GlySerArgleuAspSerTPrgluJlaleuAspSerThrsValPheqInCyasaPPro 1172
Db 4768 GGAATCGAATATGGAGTGAAGAGAACTGGGAGACATGTTGTTTCAATGTGACCA 4827
```

```

QY 1173 G1YrYAlaLeuGlnGlySerAlaGluLeuSerCysValLeuGluAsnArpHe 1192
Db 4828 GGAATGAACTTCAGAGAGAGAGAAATACCTTCAGGTGAAATTCGACTTC 4887
QY 1193 TTPGlnProSerProProThrCysIleAlaProCysGlyGlyAspLeuThrGlyProSer 1212
Db 4888 TGGCAGCCAGCCAGCCAGCTGTATAGCACCCTGGAGCAATTTAAACAGATCTTCA 4947
QY 1213 G1YAlaIleLeuSerProAntyProGluProGlyProProGlyGlyGluAspTTP 1232
Db 4948 GGCCTTATCTTTCACCAACTCTTCCTCATTCATCCGCAATAGCAGAGCTGAGACTGG 5007
QY 1233 LysValIleThrValSerProAspTyrValIleAlaLeuValPheAsnIlePheAsnLeuGlu 1252
Db 5008 ACTATCACCCGCAATAGCAGACTATGTATCTCTTGAGGTTCAATGAGTTTACATACAA 5067
QY 1253 ProGlyTyrAspPheLeuHisIleTyrAspGlyGlyAspSerLeuSerProLeuIleGly 1272
Db 5068 CCAAACTATGACTCTCTCATATCTATATGATGACACAGCACTAAATAGCCCACTGATTGGA 5127
QY 1273 SerPheTyrGlySerGlnLeuProGlyValArgIleGluSerSerAsnSerLeuPheLeu 1292
Db 5128 AGTTTCAAGACAGACAGTACCGAGAGATAGAAAGCACTCAATATACATCATTTG 5187
QY 1293 AlaPheArgSerAspAlaSerValSerAsnAlaGlyPheValIleAspTyrThrGluAsn 1312
Db 5188 GCTTTTCGAGTGAATGATCTGTATGATTACACTGATTTTCATAGAAATCAAGACAAA 5247
QY 1313 ProArgGluSerCysPheAspProGlySerIleLysAsnGlyTyrArgValGlySerAsp 1332
Db 5248 CTGCAGAGCTCTCTGTTGATCCAGGCAATATATATGATGACACCAACTGGAGATGAT 5307
QY 1333 LeuLysLeuGlySerSerValThrTyrTyrCysHisGlyGlyTyrGluValGluGlyThr 1352
Db 5308 TATTAATTAGGGTCAACAGCTACCTATATACGTGATGCTGTTATGTTCTCAAGTTAT 5367
QY 1353 SerThrLeuSerCysIleLeuGlyProAspGlyLysProValTyrAsnAsnProAspPro 1372
Db 5368 TCACACCTCACCTGATATCATAGGAGATGAGAAAGCACTGATGAGATAGACCTTGCCA 5427
QY 1373 ValCysThrAlaProCysGlyGlyGlnTyrValGlySerAspGlyValIleLeuSerPro 1392
Db 5428 AGTTGTCATGCGCCTGTGGAGATCGTTCAACAGGTTCAAGGCACTGTTCTATACCA 5487
QY 1393 AsnTyrProGlnAsnTyrThrSerGlyGlnIleCysLeuTyrPheValThrValProLys 1412
Db 5488 AACTATCCAAATAATTCAGGTGGGACATATATGTTATCTATATAGACATTCACAAAG 5547
QY 1413 AspTyrValValPheGlyGlnPheAlaPhePheHisThrAlaLeuAsnAspValValGlu 1432
Db 5548 GAGTTTGTGGTGTGGCCAGTTGATTTTTCACAGCATCACTCCACGATGTTGTTGAG 5607
QY 1433 ValHisAspGlyHisSerGlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThr 1452
Db 5608 GTGATATATGGGCCACTCAGCAATCTTCTGTATCTTCCTCCAGGATCCCACTTCA 5667
QY 1453 G1YGluSerLeuProLeuAlaThrSerAsnGlnValLeuIleLysPheSerAlaLysGly 1472
Db 5668 GGAAGATCACTTCACAGTCAAGTATCAGATACAGATCAATTCGATTTCTTCAGTTGGA 5727
QY 1473 LeuAlaProAlaArgLysPheHisPheValTyrGlnAlaValProArgThrSerAlaThr 1492
Db 5728 CCAATAACAGCTAAGGATTTCACTTGTTCACCAAGCTGTTCTAGAACAAAGTTCTACA 5787
QY 1493 GlnCysSerSerValProGluProArgTyrGlyLysArgLeuGlySerAspPheSerVal 1512
Db 5788 CAATGACAGTCTGTGCTGAACCAAGATTCGGAAGAAAGAAATTTGCAATTTGCAAGTC 5847
QY 1513 GlyAlaIleValAlaArgPheGluCysAsnSerGlyTyrAlaLeuGlnGlySerProGluIle 1532
Db 5848 GGTTCATCGGTTCTTTTATGTATTCAGAGATATATTTCTCATGAGATCCATACCAATT 5907
QY 1533 GluCysLeuProValProGlyAlaLeuAlaGlnTyrAsnValSerAlaProThrCysVal 1552

```

```

Db 5908 AGGTGGAAACAGTGGCCCAATCTTGGCCCAAGTGAATGATCTTACTTACTTAT 5967
QY 1553 ValProCysGlyGlyValAsnLeuThrGluArgArgGlyThrIleLeuSerProGlyPhePro 1572
Db 5968 GTGCCCTGTGGTGAATTTTAACTAAGCCCAAGAGCAATATTTGTCACCTGGATACCT 6027
QY 1573 GluProTyrLeuAsnSerLeuAsnLysValTyrPylValValProGluGlyAlaGly 1592
Db 6028 GAGCTTATAGCAACAATCTGAATGTGTGGAAGATCAGAGTGCAGAGGAGCTGGC 6087
QY 1593 IleGlnIleGlnValValSerPheValThrGluGlnAsnTyrAspSerLeuGluValPhe 1612
Db 6088 ATTCAAGTGAAGTTGTTAGCTTGTCTACAGAACTAAATGGGAATTTCTGCACTTTAT 6147
QY 1613 AspGlyAlaAspAsnThrValThrMetLeuGlySerPheSerGlyThrThrValProAla 1632
Db 6148 GATGGGGAGACAAATGCTCCAGACTTGGAACTATTCAGAACAAATACCCCAT 6207
QY 1633 LeuLeuAsnSerThrSerAsnGlnLeuTyrLeuHisPheTyrSerAspIleSerValSer 1652
Db 6208 CTTTGAAATGACGCTAAATATCTGATCTAAATTTTCAATCAGACATCAGTGTCT 6267
QY 1653 AlaAlaGlyPheHisIleGlnGlyTyrThrValGlyLeuSerSerCysProGluProAla 1672
Db 6268 GCTGCAGATTTCACTTGAATACACAGCAATGGTTTGATCTCTGTCGACCAACA 6327
QY 1673 ValProSerAsnGlyValLysThrGlyGluArgTyrLeuValAspAspValValSerPhe 1692
Db 6328 ACTCCTAGCAGAGTAATTAATTTGAGACAGATATATGTTGAGATGATGATCTT 6387
QY 1693 GlnCysGluProGlyTyrAlaLeuGlnGlyHisAlaHisIleSerCysMetProGlyThr 1712
Db 6388 CAGTGTGATCAAGATATCTTCTTCAGGGTCACTTCACATTAATGATATCCAGGACCT 6447
QY 1713 ValArgArgTyrAsnTyrProProProLeuCysIleAlaGlnCysGlyGlyThrValGlu 1732
Db 6448 GTAAGAGATGGAATATTCATCCAAATGTTGTTAGCTCAGTGTGGTGTGCTATGCA 6507
QY 1733 GluMetGluGlyValIleLeuSerProGlyPheProGlyAsnTyrProSerAsnMetAsp 1752
Db 6508 GACTTCAGTGTGTGATCTTACGCTCGGGTTCCTGGAACTATCCACAGCACTTATGAT 6567
QY 1753 CysSerTyrPylIleAlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAspPhe 1772
Db 6568 TGCAATGAGAAATTAATCTTACCATATAGTGTGTGTACTCTCCAGTTGTAAATTT 6627
QY 1773 SerThrGluProAsnHisAspTyrIleGluIleArgAsnGlyProTyrGluThrSerArg 1792
Db 6628 TCTACAGAAACCATACATGATATTATGGAAGTACGAAGTGAATCCACAAACTAGTACT 6687
QY 1793 MetMetGlyArgPheSerGlySerGluLeuProSerSerLeuLeuSerThrSerHisGlu 1812
Db 6688 GTTATGGCGGCTTATGGTCTCAAAATACCAATCTTCTTATTCACACACCCCAAGAA 6747
QY 1813 ThrThrValTyrPheHisSerAspHisSerGlnAsnArgProGlyPheLysLeuGlyTyr 1832
Db 6748 ACCAGCTTATATTTTTCACAGTACTATTCACAAACAAACAAAGGTTTCAATATGTATAC 6807
QY 1833 GlnAlaTyrGluLeuGlnGluCysProAspProGluProPheAlaAsnGlyIleValArg 1852
Db 6808 CAAGCTTATAGTTGCAAACTGCTGATCCAGCCCGCTTGGAAATGATTTGTATAT 6867
QY 1853 GlyAlaGlyTyrAsnValGlyGlnSerValThrPheGlnCysLeuProGlyTyrGlnLeu 1872
Db 6868 GGTATATATTTTACTGTGGGTCAAAACATTTCTTGAATGTTTCCAGAGATACACATTA 6927
QY 1873 ThrGlyHisProValLeuThrCysGlnHisGlyThrAsnArgAsnTyrAspHisProLeu 1892
Db 6928 ATTGAAATTCACACTGCATCCTTCACGGAGTCAAGTGTATATGGAATATCCACTT 6987
QY 1893 ProLysCysGluValProCysGlyGlyAsnIleThrSerSerAsnGlyThrValTyrSer 1912

```


Db 6988 CCAAGGTGTGAAGCTTTGTGGTGGGAAATATACTGCATGATGAGCACCATTATCT 7047
Qy 1913 ProGlypheProserProTySerSerSerGlnAspCyValTProLeuIleThValPro 1932
Db 7048 CCGGGGATGCTGTGATGATATCAAACTTCGAAATGTTTGGCTTGTGAAGAGTACC 7107
Qy 1933 IleGlyIleGlyValArgLeuAsnLeuSerLeuLeuGlnThGluProSerGlyAspPhe 1952
Db 7108 CTTGGGAATGCGATCTACATCATATTTTACTCTTCAACACAAACCAATATGATTTTC 7167
Qy 1953 IleThrIleTyrAspGlyProGlnGlnThrAlaProArgLeuGlyValPheThrArgSer 1972
Db 7168 ATTAGCTGATGAGTGGACGACGACCAAAATTCACCTCGATGGGTGAGTGGGCAAT 7227
Qy 1973 MetAlaIleValSerThrValGlnSerSerSerAsnGlnValLeuLeuLysPheIleArgAsp 1992
Db 7228 ACCGCTTTGGAAATAGCTACAGTACAGTCAATTCGAAATTCGAAATTCGACAGTAT 7287
Qy 1993 AlaAlaThrGlyGlyIlePheAlaIleAlaPheSerAlaTyrProLeuThrLysCyPro 2012
Db 7288 TTCACAAACAGTGGCTTTTGTGCTGCTACGCTTATCAGCCTTATCAATGAGGTGTCCAA 7347
Qy 2013 ProProThrIleLeuProAsnAlaGlyValValThrGluAsnGluGluPheAsnIleGly 2032
Db 7348 CCTCCACACCTGTGGCCCAATGCTGAATTTTGAACGGAATGATGATTTGAATAGGT 7407
Qy 2033 AspIleValArgTyrArgCyLeuProGlyPheThrLeuValGlyAsnGluIleLeuThr 2052
Db 7408 GATTTTATTTAGTATGAGTGTCTTCGAGATTTACTTATGTTGTAATGCAATTCCTACG 7467
Qy 2053 CysLeuLeuGlyThrTyrLeuGlnPheGluGlyProProIleCysGluValIleCys 2072
Db 7468 TGCAGATTTAGAGAAACGACTGACATGATGATGAGCAGCTCCAGTTGTCAAGTCTCTGT 7527
Qy 2073 ProThrAsnGluLeuLeuThrAspSerThrGlyValIleLeuSerGlnSerTyrProGly 2092
Db 7528 CCTCCCAATGAAATACGGCTGATTTCTACTGAGCATATTTGACCCCTGGATCTCGAC 7587
Qy 2093 SerTyrProGlnPheGlnThrCysSerThrLeuValArgValGluProAspTyrAsnIle 2112
Db 7588 AGTTAACCAATCTTCAAAATGTGTGCAATGAGCATTTCAAGTGGAAAGGGTTAAATATC 7647
Qy 2113 SerLeuThrValGlyTyrPheLeuSerGluLysGlnTyrAspGluPheGluIlePheAsp 2132
Db 7648 ACCATGTTTGTAGAAATCTTCCACAGCAAAAGAAATTTGATGTTCTTCAGGTGATGAT 7707
Qy 2133 GlyProSerGlyGlnSerProLeuLeuValAlaLeuSerGlyAsnTyrSerAlaProLeu 2152
Db 7708 GGACCAAAATATTCAAAGTCCAGTCTTATTTCCCTCAGTGGGATTTATTCATCTGCTTT 7767
Qy 2153 IleValThrSerSerSerSerAsnSerValTyrLeuArgTyrPheSerSerAspPheAlaTyrAsn 2172
Db 7768 AATTTAAACAGCAATGTCATGAGATTTCTTCAGTGTGCAGACAGATTCATGGCAATAC 7827
Qy 2173 ArgLysGlyPheLysIleArgTyrSerAlaProTyrCysSerLeuProArgAlaProLeu 2192
Db 7828 AAAAAGGCTTCCGATAGATATATAGCTTTCTACTGATATACCAAGAAATCCCACT 7887
Qy 2193 HisGlyPheIleLeuGlnGlnThrSerThrGlnProGlySerIleHisPheGlyCys 2212
Db 7888 CATGATATATTTATCAGTCAGACAGGTGGCGCATTAACAGTGTGTCGCTGGGCGCTGT 7947
Qy 2213 AsnAlaGlyTyrArgLeuValGlyHisSerMetAlaIleCysThrArgHisProGlnGly 2232
Db 7948 GATGAGAGATTCACACTTGTGGAAAAAGCAGTGTGTGCACAAAGCTTCTCTAATGGG 8007
Qy 2233 TyrHisLeuThrPheSerGluAlaIleProLeuCysGlnAlaLeuSerCysGlyLeuProGlu 2252
Db 8008 TATATGATGATGAGATGCGCCAGTCCCTGCTGTCAAGCAATTTCTGTGGATCTTAA 8067
Qy 2253 AlaProLysAsnGlyMetValPheGlyLysGluTyrThrValGlyThrLysAlaValTyr 2272
Db 8068 GCTCCAAACAAATGAGAAATACTAAACAACAGACTATTTGTAGGAACGCGAGTTAACTAT 8127

Qy 2273 SerCysSerGlyGlyTyrHisLeuGlnAlaGlyValAsnAlaThrAlaGluCysLeuAsp 2292
Db 8128 TTTTGATATGATGATATGATTCATGTCATCCAAAGACTCACTACAGCTGATAGCAATCA 8187
Qy 2293 ThrGlyLeuThrPheSerAsnArgAsnValProProGlnCysValProValThrCysProAsp 2312
Db 8188 GATGGAAACATGAGCAATCATTAACAAGACCCCTCGCTGTGTGTGTATGATGCCAAGC 8247
Qy 2313 ValSerSerIleSerValGluHisGlyArgTyrArgLeuIlePheGluThrGlnTyrGln 2332
Db 8248 ATCAATTCCTTTATCTTGGAAACATGGAAGATGGCGAATTTGAAATGAGCTCCATTAAGA 8307
Qy 2333 PheGlnAlaLeuMetLeuIleCysAspProGlyTyrTyrTyrThrGlnArgVal 2352
Db 8308 TACAAACCAAGATGTTTACAGTGTGACCTGTGTATGATGAGCTAGGTCTCGCTCC 8367
Qy 2353 IleArgCysGlnAlaAsnGlyLysTyrPheSerLeuGlyAspSerThrProThrCysArgIle 2372
Db 8368 ATCGAATGCTCTTCTTAATGTTACTGTGAGTGGAGAAATGAAGACCATATGCCAATT 8427
Qy 2373 IleSerCysGlyGluLeuProIleProProAsnGlyHisArgIleGlyThrLeuSerVal 2392
Db 8428 ATTTCTGTGAGAACTACCTACACCTCAAAATGAAATGAATGGAACTCAAACTTCA 8487
Qy 2393 TyrGlyAlaThrAlaIlePheSerCysAsnSerGlyTyrThrLeuValGlySerArgVal 2412
Db 8488 TATGGCTCAACAGTATCTTTACTGTGAGCTGGAGATTCAGCTTGTGGCTCTCGTGA 8547
Qy 2413 ArgGlyCysMetAlaAsnGlyLeuThrPheSerGlySerGlyValArgCysLeuAlaGlyHis 2432
Db 8548 AGGGAATGCTTTCCTCAGTCTTTGAGAGATCTGAACCGATGCCATAGCGGATCAT 8607
Qy 2433 CysGlyThrProGluProIleValAsnGlyHisIleAsnGlyLysAsnTyrSerTyrArg 2452
Db 8608 TGTGCAATTCAGAACTGATGTAATGTCAGTCAATGTCAGAAATTTGATGATAGA 8667
Qy 2453 GlySerValValTyrGlnCysAsnAlaGlyPheArgLeuIleGlyMetSerValArgIle 2472
Db 8668 GACCACTGTATATCATATTAATCTGATTTTGATGATGATGTTCTTCAAGAGGATA 8727
Qy 2473 CysGlnAlaAspHisThrPheSerGlyLysThrProPheCysValProIleThrCysGly 2492
Db 8728 TGTCAACAGATTCACATTTGCTGTGTGCTGCTCCATCCCTGTGTGCTGTAGTGTGT 8787
Qy 2493 HisProGlyAsnProValAsnGlyLeuThrGlnGlyAsnGlnPheAsnLeuAsnArgVal 2512
Db 8788 CACCTGTGATGCCAATTTATGGAAGAAAGTGAATGGTTCAACTTTAATGATGTG 8847
Qy 2513 ValLysPheValCysAsnProGlyTyrMetAlaGluGlyAlaAlaAsnSerGlnCysLeu 2532
Db 8848 GTAACATTTCAATGCAATATTTGGTATCTTATGCAAGGGCCAAACAAAGGACAGTGCAG 8907
Qy 2533 AlaSerGlyGlnThrPheSerAspMetLeuProThrCysArgIleIleAsnCysThrAspPro 2552
Db 8908 GCCAACAGAAAGTGGACCATCTCCACCTATGTGCAAAATGGTGCACATGTTCTGATCCT 8967
Qy 2553 GlyHisGlnIleAsnSerValArgGln-----ValHisAlaSerGlyProHisArg 2569
Db 8968 GGAATTCACACCAATTTCTAAAGAAAGAAAGTAAATAGAAATGAGAAAT----- 9015
Qy 2570 PheSerPheGlyThrThrValSerTyrArgCysAsnHisGlyLysPheTyrLeuLeuGlyThr 2589
Db 9016 TTTACTTACGCGACCTGTGATTTCTATGACTGCAATCTGAAATATTTTATTTGATGATCT 9075
Qy 2590 ProValLeuSerCysGlnGlyAspGlyThrTyrAspArgProArgProGlnCysLeuLeu 2609
Db 9076 TCACTTTTGAATGTCAACAAATGACAAATGGAGCAAACTTTTACAGAAATGATATGATG 9135
Qy 2610 ValSerCysGlyHisProGlySerProProHisIleSerGlnMetSerGlyAspSerTyrThr 2629
Db 9136 ATTGACTGTGACACCTCGCTGCTCTTAATGACAGTCTGTGTGCGGAGGATATACT 9195

[illegible]

Db	Accession	Source	Location/Qualifiers
Db	10276	AGTGTGTCAGCTTCATTGCAAAAAGACACCTTTCAGAGGCTTACAAACGACC	1033
Qy	3010	CysleuproAenleuthrttspseGlyThnProProApsCysValProHisHsCysArg	3029
Db	10336	TGCCTCCCTGATCTTACGTGGAGTGGGATTCACGCTGAATGACATACCCACACTGTAA	10399
Qy	3030	GlnProGlnhPrProThHsSaIaAenValGAlaLeuApsleuProserMetGlyTr	3049
Db	10396	CAGCCAGAACTCTGCTCATGCAAAATGCTGAGGATGACCTTCATCTCATGGTAA	10455
Qy	3050	Thrlleuile 3052	
Db	10456	ACACTGAT 10464	
RESULT 11			
LOCUS	AY017307		
DEFINITION	Homo sapiens CUB and sushi multiple domains protein 1 short form		
ACCESSION	AY017307		
VERSION	AY017307.1		
KEYWORDS	GI:14787180		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 10944)		
TITLE	Sun, P.C., Uppaluri, R., Schmidt, A.P., Pashia, M.E., Quant, E.C.,		
JOURNAL	Sunwoo, J.B., Gollin, S.M. and Scholnick, S.B.		
MEDLINE	Transcript map of the 8p23 putative tumor suppressor region		
PUBMED	21365705		
REFERENCE	11472063		
AUTHORS	2 (bases 1 to 10944)		
TITLE	Sun, P.C., Uppaluri, R., Schmidt, A.P., Davis, M.E., Quant, E.C.,		
JOURNAL	Sunwoo, J.B., Gollin, S.M. and Scholnick, S.B.		
MEDLINE	Direct Submision		
PUBMED	Submitted (08-JAN-2001)		
REFERENCE	School of Medicine, Box 8115, 517 S. Euclid Ave, Saint Louis, MO		
TITLE	43110, USA		
JOURNAL			
FEATURES			
source			
	1..10944		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="8"		
	/map="8p23.2"		
	286..10455		
	/note="PCSMOI short form; alternatively spliced; lacks		
	exons 58 and 59"		
	/codon_start=1		
	/product="CUB and sushi multiple domains protein 1 short		
	form"		
	/protein_id="AAG52948.1"		
	/db_xref="GI:14787181"		
	/translation="MTAMRRPQSLLLLGLVLCARLLITAKGNCGLVQSPGITE		
	SGPFGHGYPNYANCTWIIITERNRIIDSEFTPLSENFDLSYDQPOOGLNKLVL		
	SPQSPGIVSTGSLILMTFTDPAVAGSGKALYELVPSHTGNCPEIILKGVHGR		
	FNIGKIRYSCLPGLGLEHAILLFCIVSPNGSMASDPAPRCAGCGGTLRTSS		
	ISSPFPSEVENNADCTWTLAEPDPTALVFTDPEEGYDFLEISTEASTWLG		
	NMLBPVLSKMWLRHFTSDSNRRKRGFAQFVKALIELKSGVMLEPSKGSNHN		
	SVLSGQVALVSHMCLDPLGIEPENGRRGSPFRVAGNAVQSCEDNVYLQDSKSTCR		
	VTEITLAWSHDPRICRARTCGSNLRGSPVITSPNYQVEDNAACWVITTPDPVY		
	IKLAFEEELKEGYDITLVGDAGCVGDTRSVLVLTGSSVFDLIVSNOMWHLQSD		
	DSIPGSPKAVYQIEIKGGCGDPCIYAYGRKTSPLHGDTLIECPAAFEALVGRVI		
	TCOONNOMSGNKPSCFSPENFASGILLSPRYPEKYNNAACWLIISFGSRH		
	LIRNDPWERPDPFLAKNGDISDITVLTGTFSGNEVPDSQLASSGHTVLEFQSDHTT		
	GRGRNITVITFGQNECHDPLGIPNGRFRGURLFGSSVSFCHDDGFVKTQSGSITCI		
	LQDNVAVSSTVPRCEAPCGGHLTASGVLIPGWPQYKDSILACHEWILIAKGSISIK		
	MTFGRPLETVNYITLVEKRDPAASPLIGVHGTQAPQLISTGNEFWYLLFTDSS		
	SIPIGLFHYEATLESDDCLDPLGIPVNGHRRGGDPGRISTVTFQCDQVITLSDDEPLVC		
	ERNIOMHVALPSCDALCGYITQKSGTGVLSRGGDPFVNSILNTWTVFVSGTCACT		


```

QY 439 AsnIleThrPheThrThrpheArgHisAsnGluCysProAspProGlyValProValAsn 458
Db 2344 AACATCATCTTACACCAATTTGGTGAATGATGATCCATGATCTGGGCAATTCATATAAC 2403
QY 459 GlyIleValArgPheGlyAspSerIleuGluLeuGlySerSerIleSerPheLeuCysAspIu 478
Db 2404 GGACACACGTTTGGTGACAGGTTTCTACTCCGGAGACTGGTTTCTTCCATCGATGAT 2463
QY 479 GlyPheLeuGlyThrGlnGlySerGluThrIleThrCysValIleuGluGlySerVal 498
Db 2464 GGCTTGTTCACAGACCCAGAGATCGAGTTCATCTCATTACCTGCAAGAGGAGCGTG 2523
QY 499 ValTrpAsnSerAlaValIleuArgCysGluAlaProCysGlyGlyIleuThrSerPro 518
Db 2524 GTCTGAGACTCCACCGTCCCGCTGTGAAGTCCATGTGTTGACATCTACACGCTCC 2583
QY 519 SerGlyThrIleLeuSerProGlyIleThrProGlyPheIleValAspAlaLeuSerCysAla 538
Db 2584 AGCGAGTCATTTTGGCTCTGGATGGCCAGAGATATATAGAGATTTCTTACATTGTGAA 2643
QY 539 TrpValIleGluAlaGlnProGlyIleValProIleValIleThrPheAspArgPheLeuThr 558
Db 2644 TGGATTAATGTAAGCAAAACAGGCACTCTATCAAAATGACTTTTGAAGATTTCAAGACA 2703
QY 559 GluValAsnIleValArgPheThrLeuGluValArgAspGlyArgThrIleValAspIle 578
Db 2704 GAGGTCAATTTATGACCTTGGAGTGCAGATGGCCAGCCAGTGTCTCCCATCTGATC 2763
QY 579 GlyValIleValArgIleThrGlnValProGlnPheLeuIleSerThrSerAsnIleValThr 598
Db 2764 GGCAGATGCCAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 2823
QY 599 LeuLeuPheSerThrIleValArgPheThrIleValArgPheThrIleValArgPheThrIle 618
Db 2824 CTGCTGTTCACCACTGACCAAGCCAGCTCCAGATCGGCTTCTCATCTCATTGAGACT 2883
QY 619 IleThrLeuGlnSerAspHisCysLeuAspProGlyIleProValAsnGlyIleArgHis 638
Db 2884 GTGACGCTGAGTGAGATCTCTGCTGAGCCGGGCAATCCCTGTGAACGGCCATCCGCAC 2943
QY 639 GlyAsnAspPheIleValIleValIleValIleValIleValIleValIleValIleVal 658
Db 2944 GGTGAGACCTTGGCATCAGGTCCACAGTACTTTCAGCTGTGACCCGGGATCACACTA 3003
QY 659 SerArgGlyIleProLeuGluCysGluProAspPheGlnIlePheArgAlaLeuProSer 678
Db 3004 AGTGCACAGCAGCCCTGCTGTGAGAGAGAACACAGTGAACACAGCTTGCCACAG 3063
QY 679 CysGluAlaLeuCysGlyIlePheIleGlnGlySerSerGlyThrIleLeuSerProGly 698
Db 3064 TGGACACCTCTATGTGAGGCTACATCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3123
QY 699 PheProAspPheIleProAsnAsnLeuAsnCysThrIleIleGluThrSerHisGly 718
Db 3124 TTTCACAGATTTTATCAAACTCTTAACCTGACAGTGAACATTAAGTCTCATGG 3183
QY 719 LysGlyValPhePheThrPheHisIleThrPheHisLeuGluSerGlyHisAspIleLeu 738
Db 3184 AAAGAGGTTAAATGATCTTTACACCTTTCACTTGAAGGTTCCACACATATTACTG 3243
QY 739 IleThrGluAsnGlySerPheThrGlnProLeuArgIleuThrGlySerArgLeuPro 758
Db 3244 ATCAGACAGAGATGGAAGTTTTCAGAGCCGTTGCCAGGCTCACCGGGTGGGTGGCT 3303
QY 759 AlaProIleSerAlaGlyLeuIleValIleValIleValIleValIleValIleValIle 778
Db 3304 CATACAGATCAAGCAGGCTGTGTTGAAACTTCACTGAGCCAGTTCGTTTATATCAGAC 3363
QY 779 PheSerMetSerIleGluGlyPheAsnIleThrPheSerGlyIleValAspLeuProCys 798
Db 3364 TTCTCAATTTTCGACAGGCTTCAATATCACATTTTCAAGAAATATCACTGAGAGCAGT 3423
QY 799 GluGluProGluValProAlaIleValIleValIleValIleValIleValIleValIle 818

```

```

Db 3424 GATGATCTCGAGACTCCCTGCTTACCCAGCAAGAAATGTTTTCATCTTGGTGGAGAC 3483
QY 819 ThrLeuThrPheSerCysPheProGlyIleValArgLeuGluGlyThrAlaArgIleThrCys 838
Db 3484 TCTCTACAGTTTTCCTGCTTCTCGGATATCTGTTGAAAGTCCACCAAGCTTACTTCC 3543
QY 839 LeuGlyValArgArgIleuThrPheSerProLeuProArgCysValAlaGluCysGly 858
Db 3544 CTGGGTGGGGCGCGCGCTGTGTGAGATGCACCTTCCCAAGGTGTGTGGCCGAATGTGA 3603
QY 859 AsnSerValIleThrGlyThrGlnIleLeuLeuSerProAspPheProValIleThrAsn 878
Db 3604 GCAGTGTCAAAAGAAATAGAGAAATAGATAGTCTTCCAAATTTTCATCCAAATTAATGAT 3663
QY 879 AsnAsnHisGluCysIleIleThrSerIleGlnThrGlnProGlyLysGlyIleGlnLeu 898
Db 3664 AATACCATGAGTGTATCTATAAATAGAAACAGAACCGGCAAGGCAATCCACTTGA 3723
QY 899 AlaArgAlaPheGluLeuSerGluGlyAspValIleuLysValIleValAspIleValAsn 918
Db 3724 ACAGAGGCTTCAGCTGTGTAAGAGATCTTAAGATATGATGAGAAACAGT 3783
QY 919 SerAlaArgLeuLeuGlyValIlePheSerHisSerGluMetMetGlyValIleThrLeuAsn 938
Db 3784 TCTCATCGTCCACTGGGCAAGTTCATTAATAATGAATCTTGGGCTGATCTTAACAGC 3843
QY 939 ThrSerSerSerLeuThrLeuAspPheIleThrAspAlaGluAsnThrSerGlyPhe 958
Db 3844 ACATCCAAATCACCTTGTGTGATGATTCACCAACCAATGATTCGACACCAAGAGTTT 3903
QY 959 GluLeuHisPheSerSerPheGluLeuIleIleCysGlyAspProGlyIleThrProLysPhe 978
Db 3904 CAATCACTATACAGATTTGATGTGTAATAATGAGAGTCCGGGATCCCTTAACAC 3963
QY 979 GlyIleValHisAspGluGlyHisPheAlaGlySerSerValSerPheSerCysAsp 998
Db 3964 GGTATAGAGATCGAGATGAGAGGCACTTACCAACATGATCTGTATACAGTTGCAAC 4023
QY 999 ProGlyIleThrLeuArgGlySerGluLeuLeuCysLeuSerGlyGluArgThr 1018
Db 4024 CCGGGTACCGCTCATGAGAGCAACCCCTGACTGTGTGAGTGAAGAGAGAGAGAGAGAG 4083
QY 1019 TrpAspArgProLeuProThrCysValAlaGluCysGlyIleValIleValIleValIle 1038
Db 4084 TGGACAAACCACTACCTCTGCTGATAGCGAAATGTGTGATCATCATGACAGCA 4143
QY 1039 SerGlyValIleLeuSerProGlyIleThrProAlaProIleValIleValIleValIle 1058
Db 4144 TCAGAGACGAATATGTCTCCCTGGCTATCCAGCTCCGATGACAAACCTCCACTGACAC 4203
QY 1059 ThrThrIleGluAlaGluAlaGlyCysThrIleIleGluLeuHisPheLeuValPheAspThr 1078
Db 4204 TGGATTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4263
QY 1079 GluGluValHisAspValIleuArgIleTrpAspGlyProValIleGluSerGlyValIleu 1098
Db 4264 GAGATGGCTCAGCAATCTTAAGTCTGGAGACGGCCGGTGGACATGACATCTGCTG 4323
QY 1099 LysGluLeuSerGlyProAlaLeuProLysAspLeuHisSerThrPheAsnSerValIle 1118
Db 4324 AAGAGTGAAGTGTCTCCGCTTCCGAGAGACATCCACAGCACTCACTCACTCACC 4383
QY 1119 LeuGlnPheSerThrAspPhePheThrSerIleGlnIlePheAlaIleGlnPheSerVal 1138
Db 4384 CTGACGTTCCAGACGACCTTCTATAGCAAGCTGCTTCCATCCAGTCTTCCACC 4443
QY 1139 SerThrAlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerGlyAsp 1158
Db 4444 TCATTTGACGACACTGTAAAGATCCAGATATGCCCAAAATGGCACCCGCTATGAGAC 4503
QY 1159 SerTrpGluAlaGlyAspSerThrValPheGlnCysAspProGlyIleValIleGluGly 1178

```

```
Db 4504 AGCAGAGAGGCTGGAGACACCGTCATTCAGTGTGACCTGGCTATCAGCTCCAGAGA 4563
Qy 1179 SerAlaIuIleSerCyValValIleGluAsnArgPhePheProGlnProSerProPro 1198
Db 4564 CAAGCCAAATATCAGCTGTGTGACGTGAATTAACCGGTTCTTTGGCAACAGACCCCTCT 4623
Qy 1199 ThrCyAlIleAlaProGlyGlyValAspLeuThrGlyProSerGlyValIleLeuSerPro 1218
Db 4624 ACATGCACTAGCTGCTTGTGAGGGAATCTGACGGGGCCAGAGGTGTATTTTGTCCACC 4683
Qy 1219 AsnTyProGluProTyProProGlyValGlyGluCysAspTTrpValThrValSerPro 1238
Db 4684 AACTAACCCAGACCGGTATCTCTCGTGGAGGAATGTGACTGAGAGTAAAGTGAACCGG 4743
Qy 1239 AspTyValIleAlaLeuValPheAsnIlePheAsnLeuGluProGlyTyTrpAspPheLeu 1258
Db 4744 GACTTGTATCGCTTGAATATTCAAAAGTTTCAACATGAGCCACCTATGACTTCTTA 4803
Qy 1259 HisIleTyAspGlyValAspSerLeuSerProLeuIleGlySerPheTyGlySerGln 1278
Db 4804 CACATCTATGAGGGGAGAGATTCCACAGCCCTCTATTGGAGTTACACAGGGCTCTCAG 4863
Qy 1279 LeuProGlyValGlyIleGluSerSerAsnSerLeuPheLeuAlaPheArgSerAspAla 1298
Db 4864 GCCCCAGAAAGATAGAGAGTAGCGGAACAGCTGTTTCTGGCATTTCGGAGTGAATGCC 4923
Qy 1299 SerValSerAsnAlaGlyPheValIleAspTyTrpGluAsnProAlaGluSerCyPhe 1318
Db 4924 TCCGTGGCCCTTTCAGGGTTTCGCCATTGAATTTAAAGAAACACCGGAAGCTGTGTTT 4983
Qy 1319 AspProGlySerIleValAsnGlyThrArgValGlySerAspLeuValSleuGlySerSer 1338
Db 4984 GACCCAGAAATATATGATGATGAGCAAGAGTTGGAACAGACTTCAAGCTTGGCTCCACC 5043
Qy 1339 ValThrTyTrpCyValGlyGlyValGlyValGluGlyThrSerThrLeuSerCySle 1358
Db 5044 ATCACTACACAGTGTGACTGTGCTATTAAGATCTTGACCCCTCATCATCAGCTGTGTG 5103
Qy 1359 LeuGlyProAspGlyValProValITrAsnAsnProArgProValCyThrAlaProCys 1378
Db 5104 ATTGGGGCTGATGGGAAACCTCTCGGAGCCAGAGTGTGCTGCTCATATGCTCCCTGT 5163
Qy 1379 GlyGlyGlnTyValGlySerAspGlyValValLeuSerProAsnTyProGlnAsnTy 1398
Db 5164 GAGGCGCATGACAGGATCAAGAGGGATGATTATACCAACATACCCCATATATAC 5223
Qy 1399 ThrSerGlyGlnIleCySleuTyPheValThrValProValAspTyValValPheGly 1418
Db 5224 ACAGCTGTCAATATATGCTCTTATTCATCAGGTACCAAGAAATTCGTGTGTTGGA 5283
Qy 1419 GlnPheAlaPhePheIleThrAlaLeuAsnAspValValGluValHisAspGlyHisSer 1438
Db 5284 CAGTTTCCATATTCAGACAGCCCTGAATGATTGCGAGAAATATTGATGGAACCAT 5343
Qy 1439 GlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThrArgIuSerLeuProLeu 1458
Db 5344 GCACAGCCAGACTTCTCAGCTCTCTCGGGTCTCATCTGAGGGAACATTCCTCTTG 5403
Qy 1459 AlaThrSerArgGlnValLeuIleValPheSerAlaValGlyLeuAlaProAlaArgGly 1478
Db 5404 GCTACGTCAATCAATATTCGCTCGCATTCAGTCCAAAGAGCGGTGCTCGCCGGGGC 5463
Qy 1479 PheHisPheValIleGlnAlaValProArgThrSerAlaThrGlnCysSerSerValPro 1498
Db 5464 TTCCACTTGTGTTCAGAGCTGTTCTCTGACAGTGCACCCCATCAGCTGTGTGCC 5523
Qy 1499 GlnProArgTyTrpGlyValArgLeuGlySerAspPheSerValGlyAlaIleValArgPhe 1518
Db 5524 GAGGCCAGATAGCGAAGAGATAGTGTCTGAGTTTCTGCGGGCTCCATCTGTCGATTC 5583
Qy 1519 GlnCysAsnSerGlyTyTrpAlaLeuGlnGlySerProGlnIleGluCysLeuProValPro 1538
Db 5584 GAGTGAACCCCGGATATCTGCTTCAGGGTTCCAGGGCGCTCACTGCGAGTCCGTGCC 5643

Qy 1539 GlyAlaLeuAlaGlnTrpAsnValSerAlaProThrCyValValProCysGlyGlyAsn 1558
Db 5644 AACGCTTGGACAGTGGAGGAGACACATCTCCAGCTGTGTGTGATACCTTGACGTGGCAAT 5703
Qy 1559 LeuThrGluArgArgGlyThrIleLeuSerProGlyPheProGluProTyTrpLeuAsnSer 1578
Db 5704 TTCACTCAACGAAGAGTATCATCTGTCTCCCGGGGTACCTGAGGCATCGGAAACAC 5763
Qy 1579 LeuAsnCyValITrplyIleValValProGlnGlyAlaGlyIleGlnIleGlnValIva 1598
Db 5764 TTGAACGTATATGAAAGATCATATGTTACCGAGGGCTCGGAAATTCMAAGTCCAAAGAT 5823
Qy 1598 IserPheValThrGluGlnAsnTrpAspSerLeuGluValPheAspGlyAlaAspAsn 1617
Db 5824 CAGTTTGGCAGAGAGACATGGGACTTCCCTTGAAGATCCACGATGTGGGAGATGTG 5883
Qy 1618 ThrValThrMetLeuGlySerPheSerGlyThrThrValProAlaLeuLeuAsnSerThr 1637
Db 5884 ACCGACCCAGACTGGGAGACTTCTACAGCACAGTACCGGCACTGCTGAACAGTACT 5943
Qy 1638 SerAsnGlnLeuTyTrpLeuHisPheTySerAspIleSerValSerAlaAlaGlyPheHis 1657
Db 5944 TCCAACCACTTACCTGCACTTCCAGTCAATTAAGTGTGAGAGCTGCTGTGTTCCAC 6003
Qy 1658 LeuGlyTyTrpThrValGlyLeuSerSerCyPProGluProAlaValProSerAsnGly 1677
Db 6004 CTGGAATACAAACTGATGAGTCTTGCTGCAATGCGCAACAGAGCCCTCCAGCAACAGC 6063
Qy 1678 ValIleThrGlyGlnArgTyTrpLeuValAsnAspValValSerPheGlnCysGluProGly 1697
Db 6064 ATCAAAATCGGAGATCGTATCATGTGTAAGAGTGTCTCTCCATTCAGTGCAGGCCGGG 6123
Qy 1698 TyTrpAlaLeuGlnGlyHisAlaHisIleSerCyMetProGlyThrValArgGlyTrpAsn 1717
Db 6124 TACACCTGAGGGGCCCTTCCCACTTCTGTATGACAGAGGACCGTTCCTCCGTGGAAC 6183
Qy 1718 TyTrpProProLeuCyValIleAlaGlnCysGlyGlyThrValGluGluMetGluGlyVal 1737
Db 6184 TATCCGCTCCCTGTGATGATGAACTGTGAGGGAGCGTGAACACCTTGGGTGGTGTG 6243
Qy 1738 IleLeuSerProGlyPheProGlyAsnTyProSerAsnMetAspCysSerTrpValSle 1757
Db 6244 ATCTGAGGCCCGGCTTCCAGGTTCTTACCCCAACACTTAAGCTGCACCTGAGAGATC 6303
Qy 1758 AlaLeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGluProAsn 1777
Db 6304 TCATTACCCATCGGCTATGTGTGACATATTCAGTTCTGAATTTTCTACCGAAGCTAAT 6363
Qy 1778 HisAspTyTrpIleGluIleArgAsnGlyProTyGluThrSerArgMetMetGlyArgPhe 1797
Db 6364 CATGACTTCTTGAATTAATTAATGACCTTACACACCGCCCATGATTTGACAAATTT 6423
Qy 1798 SerGlySerGluLeuProSerSerLeuLeuSerThrSerHisGluThrThrValTyPhe 1817
Db 6424 AGCGGACGAGATCTCCCGGGCGCTGTGACGACAAAGCATGAACCTCATCCATCTT 6483
Qy 1818 HisSerAspHisSerGlnAsnArgProGlyPheTySleuGlnTyGlnAlaTyGlnLeu 1837
Db 6484 TATATGATCATTGCAAAACCGGCAAGATTAACTTCTTACCAAGCTATGAATTA 6543
Qy 1838 GlnGluCyPProAspProGluProPheAlaAsnGlyIleValArgGlyAlaGlyTyTrpAsn 1857
Db 6544 CAGAAGCTGCAATCAATCCACCCCATTTCAAGAAATGGATGATCAACTCGGATTAAGC 6603
Qy 1858 ValGlyGlnSerValThrPheGluCysLeuProGlyTyGlnLeuThrGlyHisProVal 1877
Db 6604 GTGGGGCAATCAGTATCTTTCAGAGTGTATCTCGGATCACTTCTTAAGGCCATCTGTC 6663
Qy 1878 LeuThrCySglHisGlyThrAsnArgAsnTrpAspHisProLeuProValCysGluVal 1897
Db 6664 CTCACCTTGACAGTGGATCAACAGAACTGAACCTGTTTCCAAAGATGATGATGCC 6723
```

QY	1098	ProCys6glValAsn1IeThrSerSerAsnGlyThrValTyrSerProGlyPheProSer	1917
Db	6724	CTTGTGTGGAGAACCGTAACTTCTTCGAACCGGACCATCTACTCCCTGGCTTCTCAT	6783
QY	1918	ProTyrSerSerSerGlnAspCysValTyrPhe1IeThrValPro1IeGlyHisGlyVal	1937
Db	6784	GAGATATCCGATCCCGAAGAGACTGATTTGGCTCATACGGTCTCCACGGGACGAGATT	6843
QY	1938	ArgLeuAsnLeuSerLeuLeuGlnThrGluProSerGlyAspPhe1IeThr1IeTyrAsp	1957
Db	6844	TAACTCAACTTCACCCCTGTACAGACGGAAGCTGTCAACAGATTACATTACGTGTGGGAC	6903
QY	1958	GlyProGlnGlnThrLysProAspLeuGlyValPheThrArgSerMetAlaLysLysThr	1977
Db	6904	GGTCCCGATAGAACTCACCCCGACTGGAGATTTCAGGGCAACACAGCCCTCGAAAGC	6963
QY	1978	ValGlnSerSerSerAsnGlnValLeuLeuLysPheHisArgAspAlaAlaThrGlyGly	1997
Db	6964	GCGTATATGCTCCACCAACCAAGCTCGCTCAAGTTCCACAGGACTTTCAATGAGAGGC	7023
QY	1998	1IePheAla1IeAlaPheSerAlaTyrProLeuThrLysCysProProThr1IeLeu	2017
Db	7024	TTCTTGTCTCTCAATTCCACGATTTCACTCAAGAAATGTCAACCTCCCGCCGGTT	7083
QY	2018	ProAsnAlaGlyValValThrGlnAsnGlnGluPheAsn1IeGlyAsp1IeValArgTyr	2037
Db	7084	CCACAGCGCAAAATGCTTACTGAGGTATGATTTCCGAATAGAGATTTTGTAAGATAC	7143
QY	2038	ArgCysLeuProGlyPheThrLeuValGlyAsnGlu1IeLeuThrCysLysLeuGlyThr	2057
Db	7144	CACTGCCACCCCGGGGATACCTTGTGTGGAGCCGACATTCTGACTTCAGACTCAGTTC	7203
QY	2058	TyrLeuGlnPheGluGlyProProProLeuCysGlnValHisCysProThrAsnGluLeu	2077
Db	7204	CAGTTGCAATTTGAGGGTCTCTCCCAACATGTGAAGCAATCCACGAAATGAAAGTC	7263
QY	2078	LeuThrAspSerThrGlyVal1IeLeuSerGlnSerTyrProGlySerTyrProGlnPhe	2097
Db	7264	CGGACTGATCATCGGGAGTCAATCTCAGTCAGAGGATCCGGGTATTTTAACTCC	7323
QY	2098	GlnThrCysSerTyrPheValArgValGluProAspTyrAsn1IeSerLeuThrValGlu	2117
Db	7324	CAGACTTGCTTGGAGATTAAGAATGGAACCAACTACCACTTACCATCTTTGGAGAC	7383
QY	2118	TyrPheLeuSerGlnLysGlnTyrAspGluPheGlu1IePheAspGlyProSerGlyGln	2137
Db	7384	ACATTTCAAAGTAAACACAGTTGATGACCTGGAATGGTTGATGGTCTTCTGGGCA	7443
QY	2138	SerProLeuLeuLysAlaLeuSerGlyAsnTyrSerAlaProLeu1IeValThrSerSer	2157
Db	7444	AGCTCTGTGCTAGATGCTTAAAGTGGAACTATCTAACAAATCAAAATTTTACAGCAG	7503
QY	2158	SerAsnSerValTyrLeuArgTyrPheSerSerAspHisAlaTyrAsnArgLysGlyPheLys	2177
Db	7504	AGTATATCGATTATATCTCCGCTGTGTCCACGTGACCATGCCACAGTAAAGAAAGATTCA	7563
QY	2178	1IeArgTyrSerAlaProTyrCysSerLeuProArgAlaProLeuHisGlyPhe1IeLeu	2197
Db	7564	ATTGCGTATGACGACTTACTGCAATTTAACCAACCCCTGTAAGATGGGGTATTTCTA	7623
QY	2198	GlyGlnThrSerThrGlnProGlyGlySer1IeHisPheGlyCysAsnAlaGlyTyrArg	2217
Db	7624	AACGGAATGACGAGCGGTTGGAAACGAAGTCAATATTTTTCGAAGCGCTGATACCGA	7683
QY	2218	LeuValGlyHisSerMetAla1IeCysThrArgHisArgProGlnGlyTyrHisLeuTyrSer	2237
Db	7684	ATGCTCCGACCAACCAATGCAACCTGTAGCAAAACCACTTGGCATGATACCAAGGGGAC	7743
QY	2238	GlnAla1IeProLeuCysGlnAlaLeuSerCysGlyLeuProGlnAlaProLysAsnGly	2257
Db	7744	TCCCTTCACGCACTCTGCCAGCGGTGTCTGTGGATATCCCAAGATCCCAAGAAACGCT	7803
QY	2258	MetValPheGlyLysGlyTyrThrValGlyThrLysAlaValTyrSerCysSerGlnGly	2277

Db	7804	TCATTATACGGGAACGAGTTCACTTTGGACAGTAAGTGTCTATGAAATGTATGAAAGCC	7867
Qy	2278	TyrHisLeuGlnIaGlyAlaGluAlaThrAlaGluCysLeuAspThrGlyLeuTrpSer	2297
Db	7864	TTTCAGCTTGAATTCACAGCCAGCAAGCAAGCCGCTGTGTCAAGAAAGATGGCGTGTGAGT	7922
Qy	2298	AsnArgAsnAlaProProGlnCysValProValThrCysProAspValSerSerIleSer	2317
Db	7924	AACACGGGGAGCCGGCCCATGTGTAAAGCCGGTCGCTTCCCGACGATGTAAGCTCAGCTC	7983
Qy	2318	ValGluHisGlyArgTrpArgLeuIlePheGluThrGlnTrpGlnIlePheGlnIaGlnLeu	2337
Db	7984	TCAGAACATGTCACTCGAAGCGCTGTTTACAGATCTTGAATGAAGTACGGTCTCAAGTA	8043
Qy	2338	MetLeuIleCysAspProGlyTyrTyrTyrTrpGlyGlnArgValIleArgCysGlnIa	2357
Db	8044	TTGCTGAGCTGACAGTCTGTGTTACTATTAGAAAGCTCGAGGCTCTCGGTCGCAAGGCC	8103
Qy	2358	AsnGlyLysTrpSerLeuGlyAspSerThrProTrpCysArgIleIleSerCysGlyLeu	2377
Db	8104	AATGGACGCTGGACACATTAAGAAATAGAGGCCAACCTGTGCA-----	8145
Qy	2378	LeuProIleProProAsnGlyHisArgIleGlyThrLeuSerValTyrGlyAlaThrAla	2397
Db	8145	-----	8145
Qy	2398	IlePheSerCysAsnSerGlyTyrThrLeuValGlySerArgValArgGluCysMetAla	2417
Db	8145	-----	8145
Qy	2418	AsnGlyLeuTrpSerGlySerGluValArgCysLeuIaGlyHisCysGlyLeuProGlu	2437
Db	8146	-----GCTGCCACTGGGTGCTCCCAAC	8169
Qy	2438	ProIleValAsnGlyHisIleAsnGlyGluAsnTyrSerTyrArgGlySerValValTyr	2457
Db	8170	CCGATTGTGAACGGTCACATTAGTGAATGAGTGGCTTCAGTTACACAGACACCGTGTGTTAC	8229
Qy	2458	GlnCysAsnAlaGlyPheArgLeuIleGlyMetSerValArgIleCysGlnIaAspHis	2477
Db	8230	CAGTGCAAATCGTGGTTCCGGCTGTGGGAATCCCTGAGAGATATGCTCGCAAGACAC	8289
Qy	2478	HisTrpSerGlyLysTrpProPheCysValProIleThrCysGlyHisProGlyAsnPro	2497
Db	8290	AAGTGTCTGGACAAACGCTGTGTGTGCTCCCATCAACATGTGTCAACCTGGAAACCT	8349
Qy	2498	ValAsnGlyLeuThrGlnIleAsnGlyPheAsnLeuAsnArgValValLysPheValCys	2517
Db	8350	GCCACGGATTCACATATGACAGAGAGTTCACCTGATATGTCGAAATTTCCACCTGC	8409
Qy	2518	AsnProGlyTyrMetAlaGlnGlyAlaAlaArgSerGlnCysLeuIaSerGlyGlnTrp	2537
Db	8410	AACACGGGCTATTTGGCTGACAGGCGGTGTCTCGAAGCCAGTGTGCGAAGAACGGCAGTGG	8469
Qy	2538	SerAspMetLysProThrCysArgIleIleAsnCysThrAspProGlyHisGlnGluAsn	2557
Db	8470	AGTAGCCCTGCGCCACGCTGCGAGTGTATAACGTGTGATCAAGCTTTGTGAAAAAT	8529
Qy	2558	SerValArgGlnValHisAlaSerGlyProHisArgPheSerPheGlyThrThrValSer	2577
Db	8530	GCCATTCGTCACGGGCAACAGAACTTCCCTGAGAGATTTTGAATGGAATGAGTATACCTCG	8589
Qy	2578	TyrArgCysAsnHisGlyPheTyrLeuLeuGlyThrProValLeuSerCysGlnGlyAsp	2597
Db	8590	TACCATTTGCAAGAAAGGATTTCACTTGCCTGGAGTCTTCAGCCCTTGACCTGTATGGCAAT	8649
Qy	2598	GlyThrTrpAspArgProArgProGlnCysLeuLeuValSerCysGlyHisProGlySer	2617
Db	8650	GGCTTAATGGGACGATCCCTGCCCAAGTCTTGGCTATATCTGTGGACACCCAGGGGCTC	8709
Qy	2618	ProProHisSerGlnMetSerGlyAspSerTyrThrValGlyAlaValArgTyrSer	2637

Db 8710 CTGCAACCCGCTCTCACTGAGAGCTGTTACTATGAGCCGCTGTCGACTACTCC 8769
 Qy 2638 Cye11eglylysaagthrlleuVal1GlyAsnSerThrArgMetCysglYleuAerGlyHis 2657
 Db 8770 TGCAGAGGGGAGCGAGCCTCTATGGCAACGACGAGAGGTGCGGAGGAGACAGTAC 8829
 Qy 2658 TrpThrGlySerLeuProHisCysSerGlyThrSerVal1GlyVal1CysglYleuAerGly 2677
 Db 8830 TGAAGCGGGGACACTGCCCTCACTGACAGAAATATCTGATTCCTGATTCGATCCGGG 8889
 Qy 2678 IleProAlaHisGlyIleArgLeuGlyAerSerPheAerProGlyThrValMetArgPhe 2697
 Db 8890 ACCCAACACATGGGTCTCGGCTTGATGACTTTAAGACAAGAGTCTTCCTCCGCTTC 8949
 Qy 2698 SerCysGlyLeuAerGlyHisVal1leuAerGlySerSerGlyLysArgThrCysglYleuAerGly 2717
 Db 8950 TCCGTGTAATGGGGACACAGCTGAGGGGCTCCCTAACGACAGGTGTTGCTCAATGGG 9009
 Qy 2718 SerTrpSerGlySerGlnProGluCysGlyVal1IleSerCysglYleuAerProGlyThrPro 2737
 Db 9010 TCATGTCAGAGCTGCAG----- 9027
 Qy 2738 SerAsnAlaArgVal1PheSerAerGlyLeuVal1PheSerSerIleVal1TrpGlu 2757
 Db 9027 ----- 9027
 Qy 2758 CysArgGluGlyTrpAlaThrGlyLeuLeuSerArgHisCysSerValAerGlyThr 2777
 Db 9027 ----- 9027
 Qy 2778 TrpThrGlySerAerProGluCysLeuVal1IleAsnCysglYleuAerProGlyIleProAla 2797
 Db 9027 ----- 9027
 Qy 2798 AsnGlyLeuAerGlyLeuAerAerPheArgTrpAsnLysThrVal1ThrTrpGlnCysVal 2817
 Db 9027 ----- 9027
 Qy 2818 ProGlyTrpMetMetGluSerHisArgVal1SerVal1LeuSerCysThrLysAerArgThr 2837
 Db 9027 ----- 9027
 Qy 2838 TrpAsnGlyThrLysProVal1CysLeuAlaLeuMetCysLysProProLeuIlePro 2857
 Db 9028 ----- 9027
 Qy 2858 AsnGlyLysVal1GlySerAerPheMetTrpGlySerSerVal1ThrTrpAlaCysLeu 2877
 Db 9073 AATGGAACAGTGAAGGAGGAGATTTCGCGTGGGGCTCCAGCATTAAGTTACAGTGCATG 9132
 Qy 2878 GluGlyTrpGlnLeuSerLeuProAlaVal1PheThrCysglYleuAerGlySerTrpThr 2897
 Db 9133 GACGGTACACGCTCTCTCACTCCGCGCATCTCTCTGTGAGGCTCCGCGGGGTGTGAA 9192
 Qy 2898 GluGluLeuProGlnCysPheProVal1PheCysglYleuAerProGlyVal1ProSerArgGly 2917
 Db 9193 GGAAGAGTACCCCGCTGCTCTCTGTGTTCTGCGGAGACCTTGCGCATCCCGCAAGGG 9252
 Qy 2918 ArgArgGluAerAerGlyPheSerTrpArgSerSerVal1SerPheSerCysHisProPro 2937
 Db 9253 CGACTTATGGGAAAAGTTTCACTATTAAGTCCGAAGTCTTCCAGTGAATTCACA 9312
 Qy 2938 LeuVal1leuVal1GlySerProArgArgPheCysGlnSerAerGlyThrTrpSerGlyThr 2957
 Db 9313 TTTTACTCGGGATCTCTCAAGAAAGTCTCGCAAGTGAACGACGTCGAGCGGAGGAGTA 9372
 Qy 2958 GlnProSerCysIleAerProThrLeuThrCysAlaAerProGlyVal1ProGlnPhe 2977
 Db 9373 CAACCCACCTGATGATCTCTCATTAACCTGCGCCAGACCTTGATGCCACACTTT 9432
 Qy 2978 G1Y1leGlnAerAerSerGlnGlyTrpGlnVal1GlySerThrVal1leuPheArgCysGln 2997
 Db 9433 GGAATACAGAACTACTCAGAGGCTATGAGTGTGAAGACAGGTTTTTTTTCAGGTGAGA 9492

Qy 2998 LysGlyTrpLeuLeuGlnGlySerThrTrpArgThrCysLeuProAerLeuThrTrpSer 3017
 Db 9493 AAGGCTACCATATTTCAAGTTTCCAGCACTCGACTCGCTTGGCAATTTTAACATGAGT 9552
 Qy 3018 GlyThrProProAerCysVal1ProHisCysArgGlnProGluThrProThrHisAla 3037
 Db 9553 GGAATACAGACGAAATTAACCTCATGCTCGAGACAGACGAAACCCCGGACACGGC 9612
 Qy 3038 AsnVal1GlyAlaLeuAerLeuProSerMetGlyTrpThrIleuIle-ThrProAlaArg 3057
 Db 9613 GATGTGAGACCATCATCTTCTACTTTCGGCTTACACCTTAATGTAACACTCCATCA 9672
 Qy 3057 GAlaSerProSerArgVal1AerProSerThrAlaProAlaArgMetAlaIleGly 3077
 Db 9673 GGCTTTTCTCGAGGGGATCTGACACAGAAATGTAAAGCAGATGAATGAGACA 9732
 Qy 3077 nAlaSerArg 3080
 Db 9733 GGAAGTCCG 9742
 RESULT 12
 AY210419 12486 bp mRNA linear PRI 31-JUN-2003
 LOCUS Homo sapiens CUB and sushi multiple domains 3 (CSMD3) mRNA,
 DEFINITION complete cds.
 ACCESSION AY210419
 VERSION AY210419.1 GI:30908444
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 12486)
 AUTHORS Lau, W.L. and Scholnick, S.B.
 TITLE Identification of two new members of the CSMD gene family
 JOURNAL Genomics 82 (3), 412-415 (2003)
 REFERENCE 2 (bases 1 to 12486)
 AUTHORS Lau, W.L. and Scholnick, S.B.
 TITLE Direct Submission
 JOURNAL Submitted (02-JAN-2003) Otolaryngology - Head & Neck Surgery,
 Washington University School of Medicine, 660 South Euclid Ave.,
 Saint Louis, MO 63110, USA
 FEATURES
 source Location/Qualifiers
 1..12486
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..12486
 /gene="CSMD3"
 142..10647
 /note="contains 14 CUB domains and 27 sushi domains"
 /codon_start=1
 /product="CUB and sushi multiple domains 3"
 /protein_id="AA034702.1"
 /db_xref="GI:30908445"

/translation="MGIKSGFTWNLVFLITVSCVKGFITCGTLKGLNGTBSPEP
 PXYRPGANCATWYIIABERNRIVQSFVLEBEYULSLYDHPHTNTRTLTGFH
 LPPVSTKVFSLRLTSDPAVAGHFKVYEELOSSCNPPVPKPVLYGTRFDVG
 DKTRYSCVGYIIDGHQOLCIANSVNTASMDPVPICRBDACGGMRSSGSISSP
 SPENVEYHNADCTWTVAERPDLSLIFDPOMEKYDYIEIGSEPPITLWLGMINP
 PTLISKNMLRLHFTVDSNHRIRYGFSAIPYOKAIDPKSKGFLTPKDNKSNKSTLN
 EGGIKTASNLCDPGEPEENKRIQSDPSLSITVQSDDEDYVLQAKSITQRIAEVF
 AASDRPVCKVKTCSNLQPGFTSPNFPQYDSNACVWITAVNTNKYIQINE
 EERDLIGYDTLIGDGEVGDPRVLYVLTGSFVDDLIVSMSSQWMLHQTDESVS
 VGRKANYKEIEKESCGPGRPLIGIREGDSNRDYLRFECOGFELIGKSIYVCEN
 SPMANIPICIFPCLSNFAPMGTIVSPDPEVGNLNTWTIISDPGRHLSFND
 FDLSEQDFLAVDGSPSPILGTGTGAEVSHILNSNITLLEFQADSMSCGRGN
 ITYNTFGHNECDPDIPIINARRGDNFQULSSISVLCERGFITQGFETICTLMDK
 VMSSGLIPKCGACGHSFAPSCVILSPGVYKQSLCENWVIEAPGHSIKITER
 FQRELNVYLEVHDGPNLLSPILGSYNGTQVPOPLFSSSNFIVLFTTDSRNNNGK
 IHVESYVNTVYSLDQIPIVHGRRYGHDFISGTVSFSQSGYRLSHBEPILCEKNM

MSHPLPTCDALGGPVRGSPGCTLLSPGVPEPYNSLCTWTVDYTHGCVOPNHTTH
 LEDHDDYLLITENSGSFTQPIALUTSDDLPPTINAGLYNPPAOLKIFISDFISYEGN
 ITRSPNLERCEDEPQIPOTGRIFNFGIGDLTFSGSSGRLESTSEIICLGGRARY
 WAPLPRCAVACGASATNNEGILLSPNPLVNNHECTVYIIOVAGAGIISARTTH
 LAQGVNLKIDYDKTKTHLGAFTGASMRGLTSTGNSQMLMFNSDPDEGDEFOFLY
 YTSFELSHCEDEPGIPQFYKISDQHPAGSTIYICNGNYLHSLIKCTGGRAM
 DYPLSCIAECGGRFKGSSGRILLSPVPPDNNLRMMNIEVDPCNIVYLSQTLAD
 TEASHDILRWMDGPPENDMLKEISGLIPGISHLANTVITIGEDTDYISKAFIAD
 FSSVATACRDPGVPMNGTRNDGREDGDFVQCDPEYELQGERSTICLOENRYEM
 QSPSPVCIAPCCGNLTSGLFSLSPNPHYPHSDCOMTITVANADYISLAFISFI
 BENYDFLYIDGPNLSPILIGSPQSKLPERIESNSNTMLAFNSDSVSTYHLEIY
 KAKLSCDFRDNINNGTRLMYDLKSTVTYVCDAGVLYGTSTICIMDDRCPM
 NALPSCFAPGSGRSTSGEGLVSPNYKPNVSGNVCYSIAKREKPVVSGRPFOT
 SLHDVEYEDGPTQOSSLSSLSGSHSESLPLSGNITRTFTSPGPTIRAKGHFY
 OAVPTSTGCSVPEPRGRIRKNERPVGSVYFDNCPGTYLIGSLAIKCEYVNSL
 AOMNDSLPTCIVPCGGLITTKGTHLSGPEPYDNINLVCWKTIVBAGIIOVVS
 FATHENMSLDFIDGDNNAFRLSGYSTTILPHLJNSNNLYANFOSDISVSAAGFR
 LEYTAIGLDSCEPEPQTPSSGIKIDRYVGDVAFQCDQSTYLSGHITCMGPVFR
 WNPPTICLAOCGAMSDPSGLILSPGPNYPSSLDCTMTINLPIGFIHLOFVNS
 TETIDYLEVSGSSESTVIGLSGPOIPSLPSTHETISVHSPSONKOGRIY
 YOAYOLSCDPDRPRNGFVIGNDPTVQCTISPCPFGTYLIGNSALTCLAGSRMN
 HPLRCEALCGNNTANNGTISPGYDEYENFODCFMVRPENGITVINTVLOTTE
 PIYPTITWMDPDONSPOIGQSGNTALLESYSTSNOLIKFHDFTTSGFVLSYHA
 YQLRVCCPPVPVPAEILTEDDEEIGDILRYOCLPFTLVNAILTCRLGKELMDG
 APVCOVLCPNBELRLDSTGVILSPYDPSYPLNOMCAMSIVEKGYNITMFEFPOT
 EKEFDVLOVDYDPNIQSPVLISLSDYSAPNLSNHEVFLQMSADHGNKKGRFIR
 YIAFYCTSPESPFGYIISOTGOLNVVMACDRGRRLVGLKGSNAVKRSYKTHAD
 APVACOLAISGCIKPAKPNGLITTDVLYNGRYVFPNDGRILSSKELTAVCOISGT
 MSNNKTRCVYISCGELPTPPPNKIGTOTSISLAFITCDLGFMLVIGSAVECLIS
 GLMSESTRCLAGHCGIPBLVNGQVIGENYGRADYVVOYCNDFILIGSVAVICOD
 HNSGJLPSCEVPSGCHGSPYIGRTSGNENFNDVTVEFISGIMQGTAKCOCAN
 RQMSHPMPCKVNVCSDPGIPANKRESKLEHNGFNTGVFVPCNDGYELPSSVLIL
 CQPNQWMDKPLECIMIDCGHVPVPAVLSGEKTYTGVHVSCTGKSLIGOSSTRT
 COLNGHSGSQPHSGDAGTCGDPGRHSGRORSEPRKSVRAACTGTGILHSE
 ERTCLANGSWTQPECKAVOCCNPTTANKYFRIDGTFSSVYSCHEGTYLISGP
 SVRQCTANGWSGTLPNCTTISGDBPEIPANGLRGDGVVGNVSMYOCGYTMEIN
 GSRIRCTINGTWSGMPTRAVTCTPPIPSNGRLNGLETFDMFGESYICSGYELIS
 PPAVLTCGNGTWSGVEVQCLPKFCGPGIPQAQKRGKSFYIOSVFSGNCFFLIV
 GSTRICOADGTWSGSSPHCIPTOTSCENPGRHSONMTGATVOSVFSGNOHCKG
 HLLOQSTRTCLPDLTWSGIOPECIPHSCGPEPTPAHANVGMDDLSSHAYTITOP
 GFFLAGGTEHRVCRSDNTWTGKVPICAGSKILVXDPRALGTFSRKLSPVDPVFRQY
 YIMKSPNPKRGQPMPLTYSSTRASGRNATLSSNMELLSGKYSQZKALMIRI
 YLIRVPAHSAVKKKEENWAMDGVSAEPDAGTYVFOGLQGDYQFGLORGLAMS
 EGSNNSQHPGHTNSSVALAILVFPFLLPAGFPYLYKORTAPKTYQYGSVHEENN
 GOAEPENMVDNAKSVEKAVRFPDLNLTCTMV"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 12486
 Score: 10384.00 Matches: 1798
 Percent Similarity: 74.62% Conservative: 492
 Best Local Similarity: 58.59% Mismatches: 659
 Query Match: 61.14% Indels: 120
 DB: 9 Gaps: 10

US-10-016-248-2 (1-3104) x AY120419 (1-12486)

QY 3 G1YAlaProProAlaLeuLeuLeu-----ProCySerLeuLeuSer 17
 DB 925 GGTTTGAGGCACTTACCATATGTTATCTCGAATGATATACCAACCATTTATCAC 984
 QY 18 Asp-----CySerAlaSerAsnGlnArgHisSerValGly 29
 DB 985 AACAAAAAAGTGGCTCAGACTGATTTTGTTCACAGACAGCAATCATCGATACCGT---GGA 1041
 QY 30 ValGlyProSerGluLeuValIleValysGlnIleGluLeuLeuSerArgIleValIleLeu 49
 DB 1042 TTTAGGTCCTCCATATCAAGTGAAGAAAAGCCCATAGATTTTAAATCTAGAGGATTTAAATTG 1101
 QY 50 MetProSerIleAspAsnSerGlnIleValIleSerValLeuThrGlnValGlyValSerGln 69
 DB 1102 TTTCCAGGGAAGAACACAGCAACAGATTCTTCTAATCTTAATAGAGAGGATTTAAACA 1161

QY 70 G1YHisMetCysProAspProGlyIleProGluArgGlyValArgLeuGlySerAsp 89
 DB 1162 GCTTCCAATTATATGCCCAGATCCAGAGAACCGAATAAGGGAAGCAATCGATCATAGT 1221
 QY 90 PheArgLeuGlyIleSerValGlnPheThrCysAsnGluGlyTyrAspLeuGlnIleSer 109
 DB 1222 TTTAGCCTTGGAATCACTGACGCTCTTGTGATGAAGATATATCTTACAGGGCGCA 1281
 QY 110 LysArgIleThrCysMetLysValSerAspMetPheAlaIleTrpSerAspPro 129
 DB 1282 AATAGCATCACTGTCAAGAGATAGCTGAAGTTTTCCTGCTTGAGTATCAAGAGCT 1341
 QY 130 ValCysArgAlaArgMetCysAspAlaHisLeuArgGlyProSerGlyIleIleThrSer 149
 DB 1342 GTGTGTAAGTGAAACGTTGGCTCTATCTTCAAGGACCAAGTGATCTTACATCT 1401
 QY 150 ProAsnPheProIleGluIleTrpAspAsnAlaHisCysValIleIleThrAlaLeu 169
 DB 1402 CCCAATCTTCGTTCCAGATATACAGCAATGACAAATGTGTGGTCAATCACACAGTG 1461
 QY 170 AsnProSerLysValIleLysLeuAlaPheGluGluPheAspLeuGlyValArgGlyTyrAsp 189
 DB 1462 AATACAAATTAAGTTATCCAGATAAATTTGAAAGAAATTTGATCTGAGATTCGCTATAGT 1521
 QY 190 ThrLeuThrValGlyAspGlyGlnAspGlyAspGlnIleValThrValLeuTyrMetSer 209
 DB 1522 ACCCTGACAAATGGGAGGAGGGGCGCAAGTTGAGATCTTAGACAGCTCCAAAGTG--- 1578
 QY 210 GlnAsnAlaCysSerAspSerProHisThrProGlySerArgIleProGlu----- 226
 DB 1579 -----CTGACTGAAACCTTTGACAGACTTGTAAGTG 1611
 QY 227 SerMetSerGlyAspIleTrpArgGlnIleValTrpThrValLeuGlnIleCysArgAspIle 246
 DB 1612 ACGATGAGTAGCAATATGCTGCTGCACTTCAAGC----- 1647
 QY 247 SerSerSerAspAlaArgSerGlySerValArgLysSerProLysTrpSerAsnAlaVal 266
 DB 1648 -----GACGAAAGCTTGGAGTCTGTTGTTCAAGTTAACTACMAA----- 1689
 QY 267 GluLeuValAlaProGlyThrGluIleGluGlnIleGlySerCysGlyAspProGlyIlePro 286
 DB 1690 -----GAAATGAGAAAGAAAGTTGTGTGATCTCTGTCGATCACCC 1728
 QY 287 AlaTyrGlyArgGluGluGlySerArgPheHisHisGlyAspThrLeuLysPheGluCys 306
 DB 1729 TTATATGGAATTAAGAGAGCGATGATTTTCTATATGCTGATGTTTAAAGTTGAATGC 1788
 QY 307 GlnProAlaPheGluLeuValGlyGlnLysAlaIleThrCysGlnIleAsnAsnGlnIleTrp 326
 DB 1789 CAGTTTGGGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1848
 QY 327 SerAlaLysLysProGlyCysValPheSerCysPhePheAsnPheThrSerProSerGly 346
 DB 1849 TCTGCAAAACATACCATCTGATCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908
 QY 347 ValIleLeuSerProAsnIleTrpProGluAspTyrGlyAsnHisLeuHisCysValIleTrpLeu 366
 DB 1909 ACGATTTCTTCTCTGATTAACCAAGAGGATATGGAATTAATTAATTAATTAATTAATTAATTAAT 1968
 QY 367 IleLeuAlaArgProGluSerArgIleHisLeuAlaPheAsnAspIleAspValGluPro 386
 DB 1969 ATAATCTGTATCCAGGAGACCGGATATCATCTTCTTCAATGATCTTGAACCTGGAATTC 2028
 QY 387 GlnPheAspPheLeuValIleLysAspGlyAlaThrAlaGluAlaProValLeuGlyThr 406
 DB 2029 CAGTTTATTTCTTCTGTTTAAAGATGAGTCTCCAGAAATCCCAATATCTTGAAC 2088
 QY 407 PheSerGlyAsnGlnLeuProSerSerIleThrSerSerGlyHisValAlaArgLeuGlu 426
 DB 2089 TTTACTGGGCTAGAGGCTCTCATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2148
 QY 427 PheGlnThrAspHisSerThrGlyLysArgGlyPheAsnIleThrPheThrThrPheArg 446


```
QY 1167 ValPheGlnCysAspProGlyThrAlaLeuGlnGlySerAlaGluIleSerCysValIys 1186
Db 4369 GTTTTCATGTGATGACCAAGATATGAACTTCAGAGGAGAAAGAAATACCTGCATTGAG 4428
QY 1187 IleGluAsnArgPheMetProGlnProSerProThrCysIleAlaProCysGlyGly 1206
Db 4429 GTGAAATATCGTAATCTTGGCAGCCCAAGCTGTATATGACACCTGTGGAGGC 4488
QY 1207 AspLeuThrGlyProSerGlyValIleLeuSerProAsnThrProGluProThrPro 1226
Db 4489 AATTTCAGAGATCTTCAGGCTTTATCTTTCACCAAACTCCCATCATCATATCCGAT 4548
QY 1227 GlyGluCysAspThrPheValThrValSerProAspThrValIleAlaLeuValPhe 1246
Db 4549 AGCAGAGACTGTGACTGAGCTATCACCGCATGACATATGTTATCTCTGGCGCTTC 4608
QY 1247 AsnIlePheAsnLeuGluProGlyThrAspPheLeuHisIleThrAspGlyArgAspSer 1266
Db 4609 ATCAGTTTATGATGAAACCAAACTATGACTTCTATATCTATGATGACCAAGCAT 4668
QY 1267 LeuSerProLeuIleGlySerPheThrGlySerGlnLeuProGlyArgIleGluSerSer 1286
Db 4669 AATACCCACTGATTTGAAAGTTTCAAGACAGCAAGTACCAAGAGAAATGAAAGCAGC 4728
QY 1287 SerAsnSerLeuPheLeuAlaPheArgSerAspAlaSerValSerAsnAlaGlyPheVal 1306
Db 4729 TCAAATACAAATGATGCTTGGCTTTTGGAGTGATGATCTGTATTCACATCGATTTTCAT 4788
QY 1307 IleAspThrThrGluAsnProArgGluSerCysPheAspProGlySerIleLysAsnGly 1326
Db 4789 CTAGAAATCAAAAGCAAACTCGAGAGATCCGCTTTATTCAGGCAATATATCAATATGCG 4848
QY 1327 ThrArgValGlySerAspLeuLysLeuGlySerSerValThrThrCysHisGlyGly 1346
Db 4849 ACCAGACTTGGAATGGAATTAATATGAGGTCAACAGTCACTATTCGTGATGCTGCT 4908
QY 1347 TyrGluValGluGlyThrSerThrLeuSerCysIleLeuGlyProAspGlyLysProVal 1366
Db 4909 TATGTTCTTCMAAGTTTATTCMAACTCACCTGTATCATGGAGATGATGGAAGACCTGGA 4968
QY 1367 TrpAsnAspProArgProValCysThrAlaProCysGlyGlyGlnThrValGlySerAsp 1386
Db 4969 TGGAAATGAGCCTGCCAAGTGTGATGATGCGCCCTGTGGAAGTGTTCMAACAGGTTAGGA 5028
QY 1387 GlyValValLeuSerProAsnThrProGlnAsnThrThrSerGlyGlnIleCysLeuThr 1406
Db 5029 GGCACGTGTTATCACCAAACTATCCAAAATTTACAGTGTGAGACATAATGTGTTAT 5088
QY 1407 PheValThrValProLysAspThrValValPheGlyGlnPheAlaPhePheHisThrAla 1426
Db 5089 TCTATACAGATTCMAAGAGGTGTGTGTGTGGCAGATGTGATTTTCCAGACATCA 5148
QY 1427 LeuAsnAspValValGluValHisAspGlyHisSerGlnHisSerArgLeuLeuSerSer 1446
Db 5149 CTCACAGATGTGTGAGGTGATGATGAGCCCACTCAGCAATCTTCTCTGTAATCTTC 5208
QY 1447 LeuSerGlySerHisThrGlyGluSerLeuProLeuAlaThrSerAsnGlnValLeuIle 1466
Db 5209 CTCTCAGAGATCCATTCAGAGATCATCTCCATGAGTTCAAGTAATCAGATCACAAT 5268
QY 1467 LysPheSerAlaLysGlyLeuAlaProAlaArgGlyPheHisPheValThrGlnAlaVal 1486
Db 5269 CGATTACTTCACTGGAACCAATACAGCTAAGGAGTTTCACTGTTTATCCAACTGTT 5328
QY 1487 ProArgThrSerAlaThrGlnCysSerSerValProGluProArgThrGlyLysArgLeu 1506
Db 5329 CCTAGAAACAGTTCTACATGCACTGTGTGCGCGTGAACCAAGATTCGGAAGAAAT 5388
QY 1507 GlySerAspPheSerValGlyAlaIleValAlaPheGluCysAsnSerGlyThrAlaLeu 1526
Db 5389 GGCATGAATTTGACGTGGTTTCATCGTTCTTTTGTATGTAAATCCAGATATATTTCTC 5448

QY 1527 GlnGlySerProGluIleGlnCysLeuProValProGlyValaLeuAlaGlnTrpAsnVal 1546
Db 5449 CATGATTCATAGCAATTTAGGTGTGAAGAGTCCCAATCTTGTGGCCAGGTGAATGAT 5508
QY 1547 SerAlaProThrCysValValProCysGlyGlyAsnLeuThrGluArgArgGlyThrIle 1566
Db 5509 TCTTATCACTTGTATTTGGCTGTGTGTGGAATTTTAACTTAACGCAAGGACATAT 5568
QY 1567 LeuSerProGlyPheProGluProThrLysAsnSerLeuAsnCysValTrpLysIleVal 1586
Db 5569 TTGTCACTGGAACCTTGAAGCTTTATGCAACAATCTAAATGTGTGTGAAGATCACA 5628
QY 1587 ValProGluGlyAlaGlyIleGlnIleGlnValValSerPheValThrGluGlnAsnTrp 1606
Db 5629 GTGCCAGAGGAGCTGAGTTCAGTGAAGCATGTTAGCTTTGCTACAGAAATATG 5688
QY 1607 AspSerLeuGluValPheAspGlyAlaAspAsnThrValThrMetLeuGlySerPheSer 1626
Db 5689 GATTCCTGACTTTTATGATGAGGAGCAACAATGCTCAAGACTTGAAGCTATTTCA 5748
QY 1627 GlyThrThrValProAlaLeuLeuAsnSerThrSerAsnGlnLeuThrLeuHisPheThr 1646
Db 5749 GGAACAACAATACCCATCTTTTGAATAGTGTATTAATCTGTATTAATTTTCA 5808
QY 1647 SerAspIleSerValSerAlaAlaGlyPheHisLeuGluThrValGlyLeuSer 1666
Db 5809 TCAGACATCAGGTTTCTGCTGACAGATTTCTTGAATACACAGCAATGTTTGAT 5868
QY 1667 SerCysProGluProAlaValProSerAsnGlyValLysThrGlyLysArgThrLeuVal 1686
Db 5869 TCTCTCTCGAACCAACAACCTCTAGAGTGAATTAATAATTGAGACAGATATATGTT 5928
QY 1687 AsnAspValValSerPheGlnCysGluProGlyThrAlaLeuGlnGlyHisIle 1706
Db 5929 GGAATGTAGATACCTTCAAGTGTATCAAGATATCTCTTCAAGGCTCACAT 5988
QY 1707 SerCysMetProGlyThrValArgArgTrpAsnThrProProLeuCysIleAlaGln 1726
Db 5989 ACATGATGCCAGAGCTGTGAAGATGAAATATCAACCCAAATTTGTTAGCTCAG 6048
QY 1727 CysGlyGlyThrValGluGluMetGlyValIleLeuSerProGlyPheProGlyAsn 1746
Db 6049 TGTGTTGTGTATGTCAGATTCAGTGTGTATCTCTGAGCTTCTGGAATC 6108
QY 1747 TyrProSerAsnMetAspCysSerThrLysIleAlaLeuProValGlyPheGlyAlaHis 1766
Db 6109 TATCCACAGATTTAGATGTGACATGACATGAATTAATACCAATAGGTTTGTGATCAT 6168
QY 1767 IleGlnPheLeuAsnPheSerThrGluProAsnHisAspThrIleGluIleArgAsnGly 1786
Db 6169 CTCAGATTTGTAATTTTCTTACAGAAACCATCATGATTAATTTGGAAGTACAAAGTGA 6228
QY 1787 ProThrGluThrSerArgMetMetGlyArgPheSerGlySerGluLeuProSerSerLeu 1806
Db 6229 TCTCAAGAACTAGTACTGTATTTGGCGGCTTAGTGCTTCAAAATPACATCTTCTTA 6288
QY 1807 LeuSerThrSerHisGluThrThrValThrPheHisSerAspHisSerGlnAsnArgPro 1826
Db 6289 TTCCAGCAACCACTGAAGACCACTTAATTTTCAAGTGAATTTTCAACAGTACATTCACAAACAA 6348
QY 1827 GlyPheLysLeuGluThrGlnAlaThrGluLeuGlnGlyCysArgProAspProGluProPhe 1846
Db 6349 GGGTTTCATATGTATACCAAGCTTACGTTGCAAGCTGCTTATATCAAGCCCGCTTT 6408
QY 1847 AlaAsnGlyIleValArgGlyValGlyThrAsnValGlyGlnSerValThrPheGluCys 1866
Db 6409 CGAAATGGTTTGTAAATGTATGATGATTTTACGTGTGCTCAAAACATTCATTTGATGT 6468
QY 1867 LeuProGlyThrGlnLeuThrGlyHisProValIleLeuThrCysGlnHisGlyThrAsnArg 1886
Db 6469 TTCCAGAGATACATTAATTTGAATTAAGCTCTTCAATGCTTTCACGAGATCATGCT 6528
QY 1887 AsnTrpAspHisProLeuProLysCysGluValProCysGlyGlyAsnIleThrSerSer 1906
```



```

Db 6529 AATTGGAATATCATTCACCTTCCAAAGGTGGAAGCTCTTGTGTGGGAATATTAACGATG 6588
Qy 1307 AaenqlyThrValTyrSerProGlyPheProSerProTyrSerSerSerGlnAaPcyVal 1926
Db 6589 AATGGCAACCATTTATTTCTCTGGGATCTTGATGAATATCCAACTTTCAGATGTTT 6648
Qy 1927 ThrLeuLeuThrValProIleGlyHisGlyValArgLeuAaenLeuSerLeuLeuGlnThr 1946
Db 6649 TGGCTGTGAAGAGTACCCCTGGGAATGGCATCTACATCAATTTTACTGTCCTTCAACA 6708
Qy 1947 GluProSerGlyAaPhePheLeuThrIleTyrAaPglProGlnGlnThrAlaProArgLeu 1966
Db 6709 GAACCAATATATGATTTTCTATTACTGTATGGATGGACACAGCAAAATTCACCTCAAGATC 6768
Qy 1967 GlyValPheThrArgSerMetAlaLeuLysThrValGlnSerSerSerAaenGlnValLeu 1986
Db 6769 GGTGAGTTCAAGTGGCAATACCGCTTTGGAAATCAAGTCTTCAATCAATCAATCTTA 6828
Qy 1987 LeuLysPheHisArgAaPheAlaAlaThrGlyLysIlePheAlaIleAlaPheSerAlaTyr 2006
Db 6829 ATCAAAATTTCCACAGTATTTTCAACAAGTGGCTTTTGTGCTCAGTTATCACGCTAT 6888
Qy 2007 ProLeuThrLysCysProProProThrIleLeuProAaenAlaGlnValAlaThrGlnAaen 2026
Db 6889 CAATTAAGGCTGCTGCCAACTCCACCAACCTGTGCCAATGCTGAATTTTGAACGGAAAT 6948
Qy 2027 GluGluPheAaenIleGlyAaPheIleValArgTyrArgCysLeuProGlyPheThrLeuVal 2046
Db 6949 GATGAATTTTGAAATAGGATATTTATAGTATCAGTGTCTTCCAGAGTTTACTTTAGTT 7008
Qy 2047 GlyAaenGluLeuLeuThrCysLeuLeuGlyThrTyrLeuGlnPheGlnGlyProPro 2066
Db 7009 GGTATGCAATTTGACGTCGACATAGAGAAACGACGACAGATGAGAGGACCTTCCA 7068
Qy 2067 IleCysGluValHisCysProThrAaenGluLeuLeuThrAaPheSerThrGlyValIleLeu 2086
Db 7069 GTTGTCAAGTCTGCTGTCTGCCAATGAATTAACGGCTAAGTTTACTCGAGTCAATTTG 7128
Qy 2087 SerGlnSerTyrProGlySerTyrProGlnPheGlnThrCysSerTyrPheValArgVal 2106
Db 7129 AGCCCTGATATCTCGACAGTTACCAATCTTCAAAATGTGATGAGATGAGATTTCACTG 7188
Qy 2107 GluProAaPyrAaenIleSerLeuThrValGlnTyrPheLeuSerGluLysGlnTyrAsp 2126
Db 7189 GAAAGGCTATATATATCATCATGTGTGTAATTTCTTCAGACAGAAAGAAATTTGAT 7248
Qy 2127 GluPheGluIlePheAaPglProSerGlyGlnSerProLeuLeuLysAlaLeuSerGly 2146
Db 7249 GTTCTTCAGCTGATGATGACCAATATTCAAAGTCCAGTCTTATTTCCCTCAGTGGG 7308
Qy 2147 AaenTyrSerAlaProLeuIleValThrSerSerSerAaenSerValTyrLeuArgTyrSer 2166
Db 7309 GATTAATTCATCTGCTTTTAATATTAACAAGCAATGTCATGAATATTTCTTCAGTGGTCA 7368
Qy 2167 SerSerHisAlaTyrAaenArgLysGlyPheLysIleArgTyrSerAlaProTyrCysSer 2186
Db 7369 GCACATATGCAATTAACAAGGCTTCGGAATGATATATAGCTTTCTACTGTAAGT 7428
Qy 2187 LeuProArgAlaProLeuHisGlyPheIleLeuGlnThrSerThrGlnProGlyGly 2206
Db 7429 ACACCAAGATCCCACTCATGATATTTATTCAGTCAGACAGGCTGGGCACTTAACAGT 7488
Qy 2207 SerIleHisPheGlyCysAaenAlaGlyTyrArgLeuValGlnHisSerMetAlaIleCys 2226
Db 7489 GTGCTCGCTGGCTGATGATGAGATTCGCACTTGTGAAAAAGAGAGCTGTGTGC 7548
Qy 2227 ThrArgHisProGlnGlyTyrHisLeuThrSerGlnAlaIleProLeuCysGlnAlaLeu 2246
Db 7549 AGAAGCTTCTCTATGGTATCATGCAATGGATGCGCCAGTCCCTGCTGCAAGCAAT 7608
Qy 2247 SerCysGlyLeuProGluAlaProLysAaenGlyMetValPheGlyLysGluTyrThrVal 2266
Db 2247 SerCysGlyLeuProGluAlaProLysAaenGlyMetValPheGlyLysGluTyrThrVal 2266

```

```

QY 2624 SerGlyAspSerThrValAlaValAlaValArgTyrSerCysIleGlyLysArgThr 2643
DB 8542 TTGGGAGAAAGTAACTTTGGGCTTACTGCTTCTGCAACGAAAGGCTCC 8601
QY 2644 LeuValGlyAsnSerThrArgMetCysGlyLeuAspGlyHisTrpThrGlySerLeuPro 2663
DB 8602 CTTTAAAGGCACTGATCAAGAACTGCGCAATGGAATGGCAATGGATGATCAACAACT 8661
QY 2664 HisCysSerGlyThrSerValGlyValCysGlyAspProGlyIleProAlaHisGlyIle 2683
DB 8662 CATTTGTCAGTGATGCTACTGGGACATGGCGATCCAGGATCTCCGGCCATGGCTCT 8721
QY 2684 ArgLeuGlyAspSerPheAspProGlyThrValMetArgPheSerCysGlyAlaGlyHis 2703
DB 8722 AGACAGAAAGCAATTTGAGAACTAAAGTACTGACTTATGCTTGATGATCGGTATC 8781
QY 2704 ValLeuArgGlySerSerGlyLysArgThrCysGlnAlaAsnGlySerTrpSerGlySerGln 2723
DB 8782 ATCTTCAATGGCTCGAAGAAAGAACATGTTAGCTAATGGCACTTGGACCGAAGGCA 8841
QY 2724 ProGlyCysGlyValIleSerCysGlyAspProGlyThrProSerAlaAsnAlaArgVal 2743
DB 8842 CCAGAGTGCAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8901
QY 2744 PheSerAspGlyLeuValPheSerSerSerIleValTyrGlyCysArgGlyLysTyr 2763
DB 8902 CGAATGATGCGACAAACATTTCTAGTCACTTATCTGCTGCTGCTGCTGCTGCTGCTGCT 8961
QY 2764 AlaThrGlyLeuLeuSerArgHisCysSerValAsnGlyThrTrpThrGlySerAspPro 2783
DB 8962 CTTTCTGACCTTCACTTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 9021
QY 2784 GlyCysLeuValIleAsnGlyAspProGlyIleProAlaAsnGlyLeuArgLeuGly 2803
DB 9022 AACTGTAATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9081
QY 2804 AsnAspPheArgGlyThrValThrValThrGlnCysValProGlyTyrMetMetGln 2823
DB 9082 GATATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9141
QY 2824 SerHisArgValSerValLeuSerCysThrValAspArgThrTrpAsnGlyThrLysPro 2843
DB 9142 TTGAATGCTCCAGAAATCAGGACCTTGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 9201
QY 2844 ValCysAlaValAlaLeuMetCysLysProProGlyLeuProAlaGlyValValGly 2863
DB 9202 ACTTGTAGAGCTGCTTACCTGCCAACTCTCCCAATCTCTAATGGAAGGCTGGAAGA 9261
QY 2864 SerAspPheMetTrpGlySerSerValThrValAlaCysLeuGlnGlyTyrGlnLeuSer 2883
DB 9262 ACAAAATTCAGAGGCTTGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9321
QY 2884 LeuProAlaValPheThrCysGlnGlyAsnGlySerTrpThrGlyGlnLeuProGlnCys 2903
DB 9322 TTCCCTGCTGCTTGTGACCTGTGAGGAATGATGCTGAGAGTGATGATACCGAGTGC 9381
QY 2904 PheProValPheCysGlyAspProGlyValProSerArgGlyLysArgGlyLysArgGly 2923
DB 9382 TTACCAAAAGTTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9441
QY 2924 PheSerThrArgSerSerValSerPheSerCysHisProProLeuValLeuValGlySer 2943
DB 9442 TTATATATACCAAGTCAAGGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9501
QY 2944 ProArgArgPheCysGlnSerAspGlyThrTrpSerGlyThrGlnProSerCysIleAsp 2963
DB 9502 AGCACAGAAATATGATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9561
QY 2964 ProThrLeuThrCysAlaAspProGlyValProGlnPheGlyIleGlnAsnAsnSer 2983
DB 9562 CCAACCAAACTCTTGTGAAACCAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9621

```

```

QY 2984 GlnGlyTyrGlnValGlySerThrValLeuPheArgCysGlnLysGlyTyrLeuLeuGln 3003
DB 9622 TTGGATTTCAAGTAAAGTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9681
QY 3004 GlySerThrThrArgThrCysLeuProAsnLeuThrTrpSerGlyThrProProAspCys 3023
DB 9682 GGGTCTCAACAGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9741
QY 3024 ValProHisHisCysArgGlnProGlyLysProThrHisAlaAsnValGlyValAlaLeuAsp 3043
DB 9742 ATACCCACAGCTGTAAACAGCCAGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9801
QY 3044 LeuProSerMetGlyTyrThrLeuIle 3052
DB 9802 CTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9828

```

```

RESULT 13
AB067481 10774 bp mRNA linear PRI 28-AUG-2003
LOCUS Homo sapiens mRNA for KIAA1894 protein, partial cds.
ACCESSION AB067481
VERSION AB067481.2 GI:34327985
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Nagase,T., Kikuno,R. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XII. The complete sequences of 60 new cDNA clones from brain which
code for large proteins
JOURNAL DNA Res. 8 (4), 179-187 (2001)
MEDLINE 21456151
PUBMED 11572464
REFERENCE 2 (bases 1 to 10774)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2001) Otsu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kitarazu, Chiba
292-0812, Japan (E-mail:cdna@f06.kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
On Aug 27, 2003 this sequence version replaced gi:15620846.
COMMENT
FEATURES
source
location/Qualifiers
1..10774
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="f08107"
/tissue_type="brain"
/note="This sequence was replaced that of fk04407 cDNA as
a representative cDNA sequence for
KIAA1894-vector:pbLuescript1 SK plus"
1..10774
/gene="KIAA1894"
<1..8934
/gene="KIAA1894"
/note="Start codon is not identified"
/codon_start=1
/product="KIAA1894 protein"
/protein_id="BAB67787.2"
/translaction="KESCDPGLPLGIREGDSNRDYLRFPCORGFELIGKSLVC
FNDPNSANIPICIFPCISNFTAPMGVTSIPYPRGXNNNICIWTIISDPSRHLIS
GNITVTNTGHNCPDPIPIKARRGDPQGLSSISVICEGFIKTQGTETITCLM
DGRVMSGLIPKGCAGCGHFSAPSGVILSPMGPGYKXDSLCEWVIEAPEPSHIT
FERFQELNVYLVLEVDGNLPLSGYNGQVQVPLFSINFIYLFTTNSRNN
GFKIHESYTVVTVSLDPIGIVHGRYVHDSIGTFSFGSDGRLRHEBPLCEK
NHWMSHPLPCALCGDVGPSGTILSGVDFYDPSNLCMTVDYTHKQVQFHF
TFPLEDHNYLLITENGSTOPLARITSGDLPPTINAGIKNPRADRIISDPSISYE
GNITRSEYNLRPCEDPGLPYQGSRIGFNFIGDITLTFSCSSGYRLGRTSITCLGGG

```

RRVASAPLPKCAECGASATNNNEGILSPNYPLVNNENNEHCISYIOVQAKGINISAR
 TPLAOGDGLKIDYKOKTHLIGAFAFGAMRGILTLSTGNCNOLMEANSITTEGDEGF
 QLYATSRELSHCEDPGIPORCYKISDQGHAGSTIITGNCPRGTHSSSLKMTGER
 RANDYPLSCIAEGGRFKGESSGILISPEPYDNMLKMMHLEIDPNVLSLOGL
 APTFEASHDILRWMDGPPENDMLKEISGLIPEGHISTNITVITIOFDIDFYISKGF
 AIOFSSSVATACDPGVPMNGTRDREBGTDFVOCDDYGEIQRDTYCTIOVERK
 YFMPSPVPCIPACGSLTGGSGFLLSPNPHYPHPCDMDMTITVADVYSLAFS
 FSIENPYDPLIYIDGPDNSPLIGSPDSDPERISSSSHMLAPSDSDSVYTGHL
 LEYKATRESCEPDGPNMTNTRGLMDYKLSYTYVYCDAYVIOGYSITLCTMGDGR
 PGNMRALPSCHAPGRKSTGSEGVLSPEVNPKNYSVGHNCVITIAPEKAEVRSIST
 QCSVPEPRGRIRIGNEFVAGSVLPDNCGYLHGSIARCTVPSIAQVMSADMSLT
 CIVPCGGILTRKGTILSPGYREYDNMLKMMHLEIDPNVLSLOGL
 LDYFDGDNNAPLRGSYSCTTILPHLSTNNLYLNOCISIASAHELYETAGID
 SCBPDPSSGICIGRVMYGDVSPDCCDGYGLOSHSTHCTMCPGRMRYPICTL
 AOCGMSDPSGVILSRPGFNGYSSIDCTWTNLPRIEGRHIOFVNFSETHDYE
 VRSGSSTSTVIGRLSGPQIPSLPSTHTSTLYFHSYDQNGOGRFIVYQVQLOSC
 PDRPRFNGFVIGNDFTVQOTISFECPPGYTLIGNSALTLCHGVSRMNPRLRCEL
 CGNITAMNGIYISPGYDPEYPODFMLVRVPGNGIYINFTVQOTEBIYDITVM
 DGPDPNSPOIGSGNTALSVSTNOILIKHSDPTTSGFVLSTHAYOLAVOCP
 PVPYNAELTEDDEFEIGDIIRYOCILPGLTGNALITGLGRLMDGAPVICOYL
 PANEHLRDSGTVILSRPDSYRLOKCMSISVEKYNITMTVEFQOTKEDEYLOV
 YDGPNTQSPVILSDYSSAFNITSNGHVFQWMSADHNNKGFIRIYATVSTP
 ESPPHGYIISQDGLQNSVVRMACDRGRFLVGSASVGRSSYGYHAMDPVACQAI
 SCGI PKAPNGGILITDYLIVGTVYFCDNGYRLSKELTAVACOSDGTWNNKTPR
 CVVYTCPSINSFIIHEGRMRIVNGSHYETKXVFSDDPGYHGPASIDCLPNGTWS
 WRNRPYCOJISGGLPTPNNGKIGTQTSYSTALPTCDLGMVLSAVRECLSGIL
 WBSERFCLAGHCGIPRELYNGQVIGENYRDTVYQCPGRILISSRITCOQDN
 WSGQLPSCVPSGCHPGSPLYGRISGNGFNPNDVYVSCNIGYLMQSPYTAOCQANQ
 WSHRPMCKVNCSDPGIPANSKRBSKIEHGNFTGVNVDNCPGRFLRGSVLTICQ
 PNCQMDKPLPECIMIDCGHPVPMNATLSEKTYFSGTIVASCGRKILGOSRPTQ
 LNHGMSGSPHSGDGTGCDGTGPGHROSBNPRKTSKTVACDGTGYILHGSER
 TCLANGSOTROPECRAVOCNGPTTANGFPRIDGTTPSSVYISGMEBYILSGEY
 ROCTANGTAGCTLPNCTIISGDPGIPANLARGDYVYVONANSYNQDPTMELNS
 RITCTINGTWSGMPCTRAVTCPTPOJISNGRLGEGVNFSGISISTISGPELSP
 AVILCVNGTWSGMPCTRAVTCPTPOJISNGRLGEGVNFSGISISTISGPELSP
 STRICQDGTGSGSSPHCIEPTQSCENPVRHSGONNVAMVMDLPSHGYTLICOPF
 LOGSTRTCLPDLTWSGIOPECI PHSCQOETPAHNVAMVMDLPSHGYTLICOPF
 FLAGCTEHRVCRDNTWTGKVPICENASGILKIDPRALTPBPKSVPRDVAONYL
 WKSYNFKGRKOPMTLTIVTSPMASTGRVNAITLNSKNELLSEGYYSKIDMLIRYIL
 IKVPDAHVSVKRMEENAMMDGFVSAEEDGATVFOGFIQKDGQFOLDGLANMSG
 SNSNPHGTNNSVAIILVPEPALIFAFGRFLYKORATPXYQYUGSVSHNNNQ
 AAFENPMYDTNAXSVEGKAVRFPDNLNTVCTMVV

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 10774
 Score: 9659.50 Matches: 1644
 Percent Similarity: 75.55% Conservative: 454
 Best Local Similarity: 59.20% Mismatches: 602
 Query Match: 56.87% Indels: 77
 DB: 9 Gaps: 3

US-10-016-248-2 (1-3104) x AB067481 (1-10774)

QY 279 SerCysGlyAspProGlyIleProAlaTyrGlyArgArgGluGlySerArgPheHis 298
 DB 7 AGTGTGCTGATCTCTGTAACACCTTATATGGAATTGAGAGGCGATGATTTCTTAAT 66
 QY 299 GlyAspThrLeuLysPheGluCysGlnProAlaPheGluLeuValGlyGlnValAla 318
 DB 67 CGTGATGTTTAAAGCTTGAATGCCAGTTTGGCTTTGAATTAATGAGAGAAATTCAT 126
 QY 319 ThrCysGlnLysValAsnAngIntTTPSerAlaLysLysPheGlyCysValPheSerCysPhe 338
 DB 127 GTTGTGACAGAAATAACCAATGCTGCAAAACATACCACCTGTATCTTCCCTGCGCTG 186
 QY 339 PheAsnPheThrSerProSerGlyValValLeuSerProAsnTyrProGluLysArgTyrGly 358
 DB 187 TCTAATCTTACTGACCAATGGGAACAGTTTCTTCTCTGATTACCCAGAAAGGTAATGGA 246
 QY 359 AsnHisLeuHisCysValTyrPheLysIleLeuAlaArgProGluSerArgTyrHisLeuAla 378
 DB 247 AATAATTTAATTCATCTGAGCATATATCTGTATCCAGGAGCCGAGTATCATCTTTCT 306

QY 379 PheAsnAspLysAspValGluProGlnPheAspPheLeuValIleLysAspGlyAlaThr 398
 DB 307 TTAAGATGACTTTTACCTTGAATCCAGTTTGAATTTCTTCTGTTAAAGTGGTACTCT 366
 QY 399 AlaGluAlaProValLeuGlyIleThrPheSerGlyAsnGlnLeuProSerSerIleThrSer 418
 DB 367 CCAGAAATCCCAATTTCTTGAACCTTTACTGAGGGGCTGAGGTGCTCCATCTTACTAGT 426
 QY 419 SerGlyHisValAlaArgLeuGluPheGlnThrAspHisSerThrGlyLysArgGlyPhe 438
 DB 427 AATAGTCACATCTGCGATTTGGAATTTACAGCTGACCACTCAATGTGAGGACGCTGCTT 486
 QY 439 AsnIleThrPheThrThrPheArgHisAsnGluCysProAspProGlyValProValAsn 458
 DB 487 AACATCACTTACACATTTTGAACATTAAGAAATGCGCTGATCTGGAATAACCAATCAAT 546
 QY 459 GlyLysArgPheGlyAspSerLeuGlnLeuGlySerSerIleSerPheLeuCysAspGlu 478
 DB 547 GCAAGGCGGTTTGGGAGCACTTCAATTAGAAAGTTCAATTTCAATTTAGTTTGAAGA 606
 QY 479 GlyPheLeuGlyIleThrGlnLysSerGluThrIleThrCysValLeuLysGluGlySerVal 498
 DB 607 GGATTTATTTAAACCCAGGAGAACAGAAACATTAACATGTATCTTATGATGAGAAAGTA 666
 QY 499 ValTyrAsnSerAlaValLeuArgCysGluAlaProCysGlyGlyHisLeuThrSerPro 518
 DB 667 ATGTGAGGTGACTGATTCCAAAATGTGAGGCCCATGTGTGTGCAATTTTACAGCTCC 726
 QY 519 SerGlyThrIleLeuSerProGlyIleTyrProGlyIlePheTyrLysAspAlaLeuSerCysAla 538
 DB 727 AGTGAAGTATTTCTTCCACAGATGCGCAGATACATCAAAAGCTTTGAATTTGAG 786
 QY 539 TyrValIleGluAlaGlnProGlyIleTyrProIleLysIleThrPheAspArgPheLysThr 558
 DB 787 TGGGTGATTTGAACCTGACACTGACACTATCAAAATTCATTTGAAATTTGACACT 846
 QY 559 GluValAspThrThrAspThrIleGluGluValArgAspGlyAlaArgThrYrserAlaProLeuIle 578
 DB 847 GAATGAAATTAATATGTTCTTGGAAAGTTCATGATGGGCCAAATCTTCTGTAACCTTGCTT 906
 QY 579 GlyValTyrHisGlyThrGlnValProGlnPheLeuIleSerThrSerAsnTyrLeuTyr 598
 DB 907 GGAATCTTACATGAGCAACCAAGTCCCAAGTTTCTATTTAGTATGATTTATATATAC 966
 QY 599 LeuLeuPheSerThrAspLysSerHisSerAspIleGlyPheGlnLeuArgTyrGluThr 618
 DB 967 CTTCTATTTTACACAGCAACAGTCTGTTCAATATATGTTTCAAAATTCATTTGAAAGT 1026
 QY 619 IleThrLeuGlnSerAspHisCysValLeuAspProGlyIleProValAsnGlyGlnArgHis 638
 DB 1027 GTTACAGTGAACAGTATTTCTTGTGACCTTGGCACTGATCTGATACATGGCCGCTAT 1086
 QY 639 GlyAsnAspPheThrValGlyAlaLeuValThrPheSerCysAspSerGlyTyrThrLeu 658
 DB 1087 GGTATGATTTCTCCATGCTGCTGATCTGTTTCAATTTAGTTGATTTAGAGATGACAGT 1146
 QY 659 SerAspGlyLysProLeuGluCysGluProAsnPheGlnTTPSerArgAlaLeuProSer 678
 DB 1147 AGTATGAAAGGCCCTTCTATGCAAAAACACATGCGGAGTCAATCCATCCCAAC 1206
 QY 679 CysGluAlaLeuCysGlyGlyPheIleGlnGlySerSerGlyIleThrIleLeuSerProGly 698
 DB 1207 TGTGATGATTAATGTGAGAGATGTTGAGGCGCTTATGGAACAATCTTATACCTGGT 1266
 QY 699 PheProAspPheTyrProAsnAsnLeuAsnCysThrTyrIleIleGluThrSerHisGly 718
 DB 1267 TACCCGGAATTTATTCATTAATTTCTGATATGTATGATGACCTGTATGTAACCAATGGA 1326
 QY 719 LysGlyValPhePheThrPheHisThrPheHisLeuGlnSerGlyHisAspTyrLeuLeu 738
 DB 1327 AAAGGTGTGAGTTCAACTTCCACACTTTTCAATTTGGAAGACATCATGACTTACTAGT 1386

```
QY 739 ILeThrgLUaNGlySerPheThrgInProLeuArgGlnLeuThrgLySerArgLeuPro 758
DB 1387 ATCAGAGAAATGGAGATTGTTTACCAACACCTGGACGCGCTGACGTTCAGATCTTCCT 1446
QY 755 AlaProIleSerAlaGlyLeuTyrgLyAsnPhenAlaGlnValArgPheIleSerAsp 778
DB 1447 CCACAAATCAATGCTGCTCTATGGAATTTCAAGGGCTCAATGCGTTTCATTTCAGAT 1506
QY 779 PheSerMetSerTyrgLyGlnPheAsnIleThrPheSerGlyTyrgAspLeuGlnProCys 798
DB 1507 TTTTCATATCATATGAAGAAATTAAACATATCTCTGAATATTAACCTTGAACCTTGT 1566
QY 799 GlnGlnProGlnValProAlaTyrgSerIleArgLyGlyLeuGlnPheGlyValGlyAsp 818
DB 1567 GAAGATCTCGCATTCCTCAATATGCTAGTGAATCGGGTTCACATTGGAGATTGGTAC 1626
QY 819 ThrLeuThrPheSerCysPheProGlyTyrgArgLeuGlnIleThrAlaArgIleThrCys 838
DB 1627 ACTGACCTCTCTCATGCTCTTCGGGTATGACACTGGAAAGAAACATCAGATCATCTGT 1686
QY 839 LeuGlyGlyArgArgLeuTrpSerSerProLeuProArgCysValAlaGlnCysGly 858
DB 1687 CTGTGGTGGCCGACGAGTGTGAGTGCACCTCTGCCAAGGTGTGTGGCTGAATGTGT 1746
QY 859 AsnSerValThrgLyThrgGlyThrLeuLeuSerProAsnPhenProValAsnTyrgAsn 878
DB 1747 GCATCTGCACGAATTAATGAAGAAATTTGCTGCTCCAAATTATCCACTCACTAATGAA 1806
QY 879 AsnAsnIleGlyCysIleTyrgSerIleGlnThrgInProGlyTyrgGlyIleGlnLeuLys 898
DB 1807 AACAAACATGAATGATTTATATGATTCAGGTTCAACAGAGAAAGGAAATCAATATTCTCA 1866
QY 899 AlaArgAlaPheGlnLeuSerGlyLysAspValLeuLysValTyrgAspGlyAsnAsn 918
DB 1867 GCCAGAACATTTCAATTAGACACAGAGAGATGCTTTAAGATTATATGAGAAAGATTA 1926
QY 919 SerIleArgLeuGlnGlyValPheSerHisSerGlnMetGlyValThrLeuAsnSer 938
DB 1927 ACGCATCATCTACTAGGTGCTTTTAACTGGGCATCTAATGCGCGAGTCACTTACTAGT 1986
QY 939 ThrSerSerLeuThrPheLeuAspPheIleThrAspAlaGlnAsnThrSerLySGlyPhe 958
DB 1987 ACTTCAATCAATCACTGCTGGTGAATTTAATTCGATCTAGAGAGACAGATGAAGGCTTT 2046
QY 959 GlnLeuHisPheSerSerPheGlyLeuLysCysGlnAspProGlyThrProLysPhe 978
DB 2047 CAACCTGTGTATCCAGTTTGAACCTGCACACTGAGAGATCTCGCATTCACAAATTT 2106
QY 979 GlyTyrgLysValHisAspGlnGlyHisPheAlaGlySerSerValSerPheSerCysAsp 998
DB 2107 GATACAAAGATCAAGTACCAAGGCACTTGTGTGTAAGCAACATCTTAATGATGCAAT 2166
QY 999 ProGlyTyrgSerLeuArgLySerGlnGlnLeuLysCysLeuSerGlyGlnArgTrp 1018
DB 2167 CCAAGGTAACCTCTCCACGGAAGTAGCGTTTCAAGTGCATGACAGAGAGAGAAAGGGA 2226
QY 1019 TrpAspArgProLeuProThrCysValAlaGlnCysGlyTyrgValArgLyGlnVal 1038
DB 2227 TGGGACATGCTCTGCTCTCTGATGTGCTGATGTGAGAGTCTGTTTAAAGAAATCA 2286
QY 1039 SerGlyGlnValLeuSerProGlyTyrgProAlaProGlyGlnHisAsnLeuAsnGlyIle 1058
DB 2287 TCAGGAAGAATCTTATCTCTGCTGCTCTTTCATATGACAAATTAACCTGGCTGGACG 2346
QY 1059 ThrThrIleGlnAlaGlnAlaGlyCysThrIleGlyLeuHisPheLeuValPheAspTrp 1078
DB 2347 TGGATGATTGAGGTAGTCTGGAATATGTTCAAGCTTCAGTTCTTCTTTGATACG 2406
QY 1079 GlnGlnValHisAspValLeuArgLyIleTrpAspGlyProValGlnSerGlyValLeuLeu 1098
DB 2407 GAAGCATCAATGATATCTGCGAGTCTGGAGCGGTCCACAGAAAGAAATATGCTTTTAA 2466
QY 1099 LysGlnLeuSerGlyProAlaLeuProLysAspLeuHisSerThrPheAsnSerValVal 1118
DB 2467 AAGGAATATGATGATCTCTTATCTCTGAAGAAATTCATAGACCTCATATATGTAAC 2526
QY 1119 LeuGlnPheSerThrAspPheThrSerLySGlnGlyPheAlaIleGlnPheSerVal 1138
DB 2527 ATCCAGTTTGACAGGATTTTATATATACCAATCTGATTTGCAATTCAGTTTCAACT 2586
QY 1139 SerThrAlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerGlyAsp 1158
DB 2587 TCTGTTGCCACGCGGTGTGTGACCCAGGGGTCCCATGATAGGACTGGAATGGAGAT 2646
QY 1159 SerTrpGlnAlaGlyAspSerThrValPheGlnCysAspProGlyTyrgAlaLeuGlnGly 1178
DB 2647 GGAAGAAACCTGGGACACTGTGTTTTCATGTGACCCAGATATGAATTCAGGA 2706
QY 1179 SerAlaGlnIleSerCysValLysIleGlnAsnArgPheThrgInProSerProPro 1198
DB 2707 GAGGAAGAATTAACCTGATTCAGTTAGGTTGAATGCGTACTTGTGACCCGACCCACCA 2766
QY 1199 ThrCysIleAlaProCysGlyGlyAspLeuThrgLyProSerGlyValIleLeuSerPro 1218
DB 2767 GTCTGATAGCACCTGTTGAGGCAATTTAACAGATCTTCAGGCTTATCTTACACA 2826
QY 1219 AsnTyrgProGlnProTyrgProProGlyTyrgGlyCysAspTrpLysValThrValSerPro 1238
DB 2827 AACCTTCCTCATCCATATCCGATAGCAGAGACTGTGACTGACTATACCGTCAATGCA 2886
QY 1239 AspTyrgValIleAlaLeuValPheAsnIlePheAsnLeuGlnProGlyTyrgAspPheLeu 1258
DB 2887 GACTATGTTATCTCTGCGCTTATCATGCTTTTGAATAGAACCAACTATGACTTCTTC 2946
QY 1259 HisIleTyrgAspGlyArgAspSerLeuSerProLeuIleGlySerPheTyrgLySerGln 1278
DB 2947 TATATCTATGATGAGCCAGACAGATTAAGCCACTGATTTGAGAGTTTCAAGACAGCAAG 3006
QY 1279 LeuProGlyArgIleGlySerSerAsnSerLeuPheLeuAlaPheArgSerAspAla 1298
DB 3007 TTACAGAGAAATTAAGAAAGACGCTCAATATCAATGCTTTGCGTCTTTCGAGAGTAAGA 3066
QY 1299 SerValSerAsnAlaGlyPheValIleAspTyrgThrgLysAsnProArgLySerCysPhe 1318
DB 3067 TCTGTATGTTCACTGATTTTCATCTGAATTAACAAGCAAAATCGAGAGTCTGCTTT 3126
QY 1319 AspProGlySerIleLysAsnGlyThrArgValGlySerProLeuLysLeuGlySerSer 1338
DB 3127 GATCCAGCAATTAATGAATGACACAGACTTGAATGATTAATAATTAGGCTCAACA 3186
QY 1339 ValThrTyrgTyrgCysHisGlyGlyTyrgGlnAlaGlnIleThrSerThrLeuSerCysIle 1358
DB 3187 GTCACTATTAAGTGTGATGCTGGTTATGTTCTTCAAGTTATTCACACTCAGCTATAC 3246
QY 1359 LeuGlyProAspGlyLysProValTrpAsnAsnProArgProValCysThrAlaProCys 1378
DB 3247 ATGGGAATATATGAACCTGATGGAATGAAGCCTTGGCAAGTTGTCACTGCGCCTGT 3306
QY 1379 GlyGlyGlnTyrgValGlySerAspGlyValValLeuSerProAsnTyrgProGlnAsnTyrg 1398
DB 3307 GGAAGCTTTCACAGGTTTCAAGAGCAGCTTCTATCACCACCAATCTTCCAAAATAATAC 3366
QY 1399 ThrSerGlyGlnIleCysLeuTyrgPheValThrValProLysAspTyrgValValPheGly 1418
DB 3367 AGGTGGACATATATGTTTATTTCTATAGCAGTTCCAAAGGAGTTT 3414
QY 1419 GlnPheAlaPheHisThrAlaLeuAsnAspValValGlnValHisAspGlyHisSer 1438
DB 3414 3414
QY 1439 GlnHisSerArgLeuLeuSerSerLeuSerGlySerHisThrgLyGlnLeuSerProLeu 1458
DB 3414 3414
QY 1459 AlaThrSerAsnGlnValLeuIleLysPheSerAlaLysGlyLeuAlaProAlaArgGly 1478
```

Db 3414 ----- 3414
Qy 1479 PheH1aPheVal1YrGlnAlaValProArgThSerAlaThrGlnCysSerSerValPro 1498
Db 3415 ----- 3415
Qy 1499 GluProArgTyrGlyLysArgLeuGlySerAspPheSerValGlyAlaIleValArgPhe 1518
Db 3457 GAAACAAGATTCCGAAGAAGAAATTGGCAATGAATTTCAGTCGGTTCATCGCTTTCTTTT 3516
Qy 1519 GluCysAnserGlyTyrAlaLeuGlnGlySerProGluIleGluCysLeuProValPro 1538
Db 3517 GATTGTAATCCAGAGATATATCTTCATGATCCATGCAATTAAGTGTGAACAGTCC 3576
Qy 1539 GlyAlaLeuAlaGlnTTPaenValSerAlaProThrCysValValProCysGlyValAsn 1558
Db 3577 AATTCTTGGCCCGAGTGAATGATTCCTTACTGATGTAATGGCCCTGGGGGAGATT 3636
Qy 1559 LeuThrGluArgArgGlyThrIleLeuSerProGlyPheProGluProCysTyrLeuAsnSer 1578
Db 3637 TTAACTAAGCGCAAGGAGCATTTTGTCTACCTGATACCTGAGCTTATGACAAACAT 3696
Qy 1579 LeuAsnCysValTyrPylsIleValValProGluGlyAlaGlyIleGlnIleGlnValVal 1598
Db 3697 CTGAATGTGTGTGGAAGATCAGTCGCAAGGAGGAGCTGGCATTCAGTGAAGTGTGT 3756
Qy 1599 SerPheValThrGluGlnAenTTPaSerSerLeuGluValPheAspGlyAlaAspAenThr 1618
Db 3757 AGCTTGTCTACAGAAACATTAATGGGATTCCTGAGACTTTAATGATGGGGAGAACACAT 3816
Qy 1619 ValThrMetLeuGlySerPheSerGlyThrThrValProAlaLeuAsnSerThrSer 1638
Db 3817 GCTCCAGAGACTTGGAAAGCTATTCAGGAACAACATACCCATCTTTGAATAGACGTCT 3876
Qy 1639 AsnGlnLeuTyrLeuHisPheTyrSerAspIleSerValSerAlaAlaGlyPheHisLeu 1658
Db 3877 AATATCTGATCTAATTAATTTTCATCAGACATCAGTGTCTGCTGAGATTTTCATCTT 3936
Qy 1659 GluTyrIleThrValGlyLeuSerSerCysProGluProAlaValProSerAnsglyVal 1678
Db 3937 GAATACACAGCAATTTGGTTGGATTCCTGCTGCAACACCAACTCTTACAGAGGATTT 3996
Qy 1679 LysThrGlyGluArgTyrLeuValAsnAspValValSerPheGlnCysGluProGlyTyr 1698
Db 3997 AAAATTTGAGACAGATATATGTTGGAGATGATATCTTTCACTGTGATCAAGATAT 4056
Qy 1699 AlaLeuGlnIleHisAlaHisIleSerCysMetProGlyThrValArgTyrAsnTyr 1718
Db 4057 TCTCTTCAAGGTCTCTCAGATTAATGATGCAAGACCTGTAAAGAATGGAATTTAT 4116
Qy 1719 ProProProLeuCysIleAlaGlnCysGlyGlyThrValGluGluMetGluGlyValIle 1738
Db 4117 CCATCCCAATTTGTTAGCTCAGTGTGTGTCTGTCTGACAGCTTCAAGTGTGTATCT 4176
Qy 1739 LeuSerProGlyPheProGlyAsnTyrProSerAnsnAspCysSerTyrPylsIleAla 1758
Db 4177 CTGAGCTCTGAGGTTCTTGAAGATATCCAGAGTTAGTTGACACATGAGCAATAAAT 4236
Qy 1759 LeuProValGlyPheGlyAlaHisIleGlnPheLeuAsnPheSerThrGluProAsnHis 1778
Db 4237 CTACCCATAGGTTTGGTGTATCATCTCAGTTTGTAAATTTTCTACGAAGAATCATCAT 4296
Qy 1779 AspTyrIleGluIleArgAnsglyProTyrGluThrSerArgMetMetGlyArgPheSer 1798
Db 4297 GATTATTTGGAAGTACAGAAAGATCTCTCAAGAACTGTACTGTATTTGGCCGGCTTAGT 4356
Qy 1799 GlySerGluLeuProSerSerSerLeuSerThrSerHisGluThrThrValTyrPheHis 1818
Db 4357 GGTCTCAAAATCATCTTCTTATTCAGCACCCATGAAACCAAGCTTATATTTTTCAC 4416
Qy 1819 SerAnsnHisSerGlnAsnArgProGlyPheLeuLeuGluTyrGlnAlaTyrGluLeuGln 1838
Db 4417 AGTACTATTCACAAACAAACAGGGTTTCATATGTATATCAAGCTTATCATGTTCAAA 4476

Qy 1839 GluCysProAspProGluProPheAlaAsnGlyIleValArgGlyAlaGlyTyrAsnVal 1858
Db 4477 AGCTGTCTGATCAGCCCGGTTCCGAATGTTTGTATTTGTAATGATTTTACTGTG 4536
Qy 1859 GlyGlnSerValThrPheGluCysLeuProGlyTyrGlnLeuThrGlnHisProValLeu 1878
Db 4537 GGTCAACATTTTCATTTGTAATGTTTCCAGATATCAGATTAATTTGGAATTCAGCTCTC 4596
Qy 1879 ThrCysGlnHisGlyThrAsnArgAsnTTPaSerHisProLeuProCysGlyValPro 1898
Db 4597 ACATGCTTTCACGAGACAGTCTTAATTTGAATCATCTCCATCAAGTGTGAAGCTTT 4656
Qy 1899 CysGlyGlyAsnIleThrPheSerAnsglyThrValTyrSerProGlyPheProSerPro 1918
Db 4657 TGTGTGGGAATATTACTGCAATGAATGAGCAATTAATTCCTGTGGATCTGATGAA 4716
Qy 1919 TyrSerSerSerGlnAspCysValTyrPylsIleThrValProIleGlnHisGlyValArg 1938
Db 4717 TATCCAAACTTTCAGATGTTGTTTGGCTTGTAAAGATACCCCTGGAGATGCACTAC 4776
Qy 1939 LeuAsnLeuSerLeuLeuGlnThrGluProSerGlyAspPheIleThrIleTTPaSerGly 1958
Db 4777 ATCAATTTTACTGTCTTCAACAGAACAAATATATCATTTTCACTGTATGGAGATGA 4836
Qy 1959 ProGlnGlnThrAlaProArgLeuGlyValPheThrArgSerMetAlaLysThrVal 1978
Db 4837 CCAGACCAAAATTCACCTCAGATGCTGATTCAGTGCAGCAATACCGTTTGAATTCAGTC 4896
Qy 1979 GlnSerSerSerAnsglnValLeuLeuLysPheHisArgAspAlaAlaThrGlyGlyIle 1998
Db 4897 TACAGATCTTCAATCAGATCTTATCAATTCACAGTATTTCAACAAGTGGCTTT 4956
Qy 1999 PheAlaIleAlaPheSerAlaTyrProLeuThrLysCysProProThrIleLeuPro 2018
Db 4957 TTTGTGCTCAGTATATCAGCTTATCACTAAGGAGTGTGCAACCTCCACCACTGTGCCC 5016
Qy 2019 AsnAlaGlyValAlaThrGluAsnGluGluPheAsnIleGlyAspIleValArgTyrArg 2038
Db 5017 AATGCTGAATTTTGAACGGAAGATGAATTTGAATGATATATTTATGATATCAG 5076
Qy 2039 CysLeuProGlyPheThrLeuValGlyAsnGluIleLeuThrCysLysLeuGlyTyrThr 2058
Db 5077 TGTCTTCCAGATTTTACTTATGTTGTAATGCAATTCAGTGCAGATTAAGGAAGCA 5136
Qy 2059 LeuGlnPheGluGlyProProPylsIleCysGlyValHisCysProThrAnsglyLeu 2078
Db 5137 CTCAGATGATGAGAGACCTCCAGTTTGTCAAGTGTCTGTCTGCTCCATGAAATTAACG 5196
Qy 2079 ThrAspSerThrGlyValIleLeuSerGlnSerTyrProGlySerTyrProGlnPheGln 2098
Db 5197 CTAGATTTCTAATGAGCATATTTGAGCTGTGATATCTGACAGTTAACCAAAATCTTCAA 5256
Qy 2099 ThrCysSerThrLeuValArgValGluProAspTyrAsnIleSerLeuThrValGluTyr 2118
Db 5257 ATGTGTGACAGACATTTTACGTGAAAGGTTATATATTCACAGATTTGTATGAAATTC 5316
Qy 2119 PheLeuSerGluLysGlnTyrAspGluPheGluIlePheAspGlyProSerGlyGlnSer 2138
Db 5317 TTCACAGCAAGAAAGAAATTTGATGTTCTTCAAGTGTATGATGACCAAAATATTTCAAAGT 5376
Qy 2139 ProLeuLeuValAlaLeuSerGlyAsnTyrSerAlaProLeuIleValThrSerSerSer 2158
Db 5377 CCAAGTCTTATTTTCCATGAGTGGAGATTTATCTGTCTTTAATATAACAGCAATAGT 5436
Qy 2159 AsnSerValTyrLeuArgTyrPserSerAspHisAlaTyrAsnArgLysGlyPheLysIle 2178
Db 5437 CATGAAGTATTTCTTCACTGTGTGACAGATCATGGAATTAACAAAAGGCTTCCGAGATA 5496
Qy 2179 ArgTyrSerAlaProTyrCysSerLeuProArgAlaProLeuHisGlyPheIleLeuGly 2198
Db 5497 AGATATATAGCTTTTCTACTGATATGATACCAAGAAATCCCATGATGATATATATACGT 5556

```

QY 2199 GlnThrSerThrGlnProGlyGlySerIleHisPheGlyCysAsnAlaGlyThrArgLeu 2218
DB 5557 CAGACAGGAGGCGCAGCTTAACAGTGTGCTTGGCTGTGATCGAGGATTCGACCTT 5616
QY 2219 ValGlyHisSerMetAlaIleCysThrArgHisProGlnGlyThrHisLeuTrpSerGlu 2238
DB 5617 GTTGGAAAAAGCAGTGTGTGTGCAGAAAGCTTCTCATGGGTATCATGATGGAGATCGC 5676
QY 2239 AlaIleProLeuGlyCysGlnAlaLeuSerCysGlyLeuProGluAlaProValAsnGlyMet 2258
DB 5677 CCAGTCCCTGCTGCTGCAGAGCAATTTCTGTGGGATTCCTAAAGCTCCCAAGAGGAGGA 5736
QY 2259 ValPheGlyGlyGlyThrValGlyThrValAlaValTyrSerCysSerGluGlyTyr 2278
DB 5737 ATATCAACACACAGACTATTTGGTAGAAGCGAGTTACTTATTTTGAATGATGATAT 5796
QY 2279 HisLeuGlnAlaGlyAlaGlyAlaThrAlaGlyCysLeuAspThrGlyLeuTrpSerAsn 2298
DB 5797 CGATTGTATCATCAAGAACTCACTACAGCTGTATGCCAATCAGATGGAACTAGAGCAAT 5856
QY 2299 ArgAsnValProGlnCysValProValThrCysProAspValSerSerIleSerVal 2318
DB 5857 CATAACAGACCCCTCGCTGTGTGTGTGTGTATCATGTCCAGACATCAATCTCTTATCTTG 5916
QY 2319 GluHisGlyArgTrpArgLeuIlePheGlyThrGlnTyrGlnPheGlnAlaGlnLeuMet 2338
DB 5917 GAACATGGAGATGGCGAATGTGAATGGCTCCATTATGAATACAAACCAAGTAGT 5976
QY 2339 LeuIleCysAspProGlyTyrTyrTyrThrGlyGlnArgValIleAspCysGlnAlaAsn 2358
DB 5977 TTCAGCTGTGACCTGTGTATCATGACATGAGTCTCTCCATCGAATGATCTTCTCAAT 6036
QY 2359 GlyTyrTrpSerLeuGlyAspSerThrProThrCysArgIleIleSerCysGlyGluLeu 2378
DB 6037 GGATCTTGAGATTGGAAGAAATGAAGAACCAATATGCCAATTTTCTCTGGAGAACTA 6096
QY 2379 ProIleProAsnGlyHisArgIleGlyThrLeuSerValTyrGlyAlaThrAlaIle 2398
DB 6097 CCTACACCTTCCAAATGGAAATTAAGATGGAACTCAACATCATATGGCTCAACAGCTATC 6156
QY 2399 PheSerCysAsnSerGlyTyrThrLeuValGlySerArgValArgGlyCysMetAlaAsn 2418
DB 6157 TTTAACCTGCAGCTTGGGATTCATGCTTGTGGCTCTGCTTAAAGGAAATGCCCTTCTCA 6216
QY 2419 GlyLeuTrpSerGlySerGlyValArgCysLeuAlaGlyHisCysGlyThrProGluPro 2438
DB 6217 GGTCTTTGAGATGATCTGAACACAGATGCCCTAGCGGCTCATTTGTGAATCCAGAACG 6276
QY 2439 IleValAsnGlyHisIleAsnGlyGluAsnTyrSerTyrArgIleGlySerValValTyrGln 2458
DB 6277 ATTGTGAATGATCAAGTCATGGAGAAATTAATGATATAGAGACACAGTTGTATATCA 6336
QY 2459 CysAsnAlaGlyPheArgLeuIleGlyMetSerValArgIleCysGlnGlnAspHisHis 2478
DB 6337 TGTATCTCGTGTTCGATTTGATTTGGTTCTTCAAGTAAATATGTAACAGATCAAT 6396
QY 2479 TrpSerGlyLysThrProPheCysValProIleThrCysGlyHisProGlyAsnProVal 2498
DB 6397 TGGTCTGGTCAAGCTCCATCTCTGTGCTGTAGCTGTGTCACCTGGTAGTCCAAT 6456
QY 2499 AsnGlyLeuThrGlnGlyAsnGlnPheAsnLeuAspValValLysPheValCysAsn 2518
DB 6457 TATGGAAGAACAGTGAATGGTTCACAACTTAATGATGGTAAACATTTCCATGCAAT 6516
QY 2519 ProGlyTyrMetAlaGlyAlaAlaArgSerGlnCysLeuAlaSerGlyGlnTrpSer 2538
DB 6517 ATTTGGTATCTTATGCAAGGGCCAAAGGACAGTGCAGGCGCAACGAGAGTGGAGC 6576
QY 2539 AspMetLeuProThrCysArgIleIleAsnGlyThrAspProGlyHisGlnGluAsnSer 2558
DB 6577 CATCTTCAACCTATGTGCAAGAGTGTCAACTGTTCTGATCTGTGAATTCAGACCAATCT 6636
QY 2559 ValArgGln-----ValHisAlaSerGlyProHisArgPheSerPheGlyThrThr 2575

```

```

DB 6637 AAAAGGAAGATGAATAATAGAACCTGGAAT-----TTTACTTACCGCAGCTGTG 6684
QY 2576 ValSerTyrArgCysAsnHisGlyPheTyrLeuLeuGlyThrProValLeuSerCysGln 2595
DB 6685 GTATCTTAATGACCTGCAATCTGTGATATTTTATTTGGATCTTCAGTTTGTATATGTCA 6744
QY 2596 GlyAspGlyThrTrpAspArgProArgProGlnCysLeuLeuValSerCysGlyHisPro 2615
DB 6745 CCAATAGCAATGGAGCAAAACCTTTACAGATGATATGATGATGACGTGGACACCTT 6804
QY 2616 GlySerProGlnHisSerGlnIleSerGlyAspSerTyrThrValGlyAlaValAlaArg 2635
DB 6805 GGGTCTCTCTTAATGACAGTCTGTGTGGAGAAATATCTTTGGGTCTACTGTCAAC 6864
QY 2636 TyrSerCysIleGlyLysArgThrLeuValGlyAsnSerThrArgMetCysGlyLeuAsp 2655
DB 6865 TATTCCTGACAGGAAGAGCTTCCCTTTAGGCGAGTATCAAGAACCTGCCAATTTGAAT 6924
QY 2656 GlyHisTrpThrGlySerLeuProHisCysSerGlyThrSerValGlyValCysGlyAsp 2675
DB 6925 GGCCATGGAGAGATCAACMACCTATGTCAGGTGATGCTACGGGACATGTGGCAT 6984
QY 2676 ProGlyIleProAlaHisGlyLysArgLeuGlyAspSerPheAspProGlyThrValMet 2695
DB 6985 CCAAGTACTCCGGCGCATGGCTCTGAACAGAAAGCAATTCAGAACTAAAGTACTGA 7044
QY 2696 ArgPheSerCysGlnAlaGlyHisIleValLeuArgGlySerSerGlyLysArgThrCysGlnAla 2715
DB 7045 CGTTATGCTTGTGAATCGTGTATCATCTTCACTGCTCAGAAAGAAACATGTTTACT 7104
QY 2716 AsnGlySerTrpSerGlySerGlnProGlyCysGlyValIleSerCysGlyAsnProGly 2735
DB 7105 AATGGCAATGGACCGGAAGGACGACAGAGCAAGCTGTGCGAGTGTGTAACCCAGGA 7164
QY 2736 ThrProSerAsnAlaArgValValPheSerAspGlyLeuValPheSerSerSerIleVal 2755
DB 7165 ACCACAGCAATGGGAAGTAGTCTCCGAATGTATGAGCAACAATTTCTAGTCACTCTT 7224
QY 2756 TyrGlyCysArgGlyGlyTyrTyrAlaThrGlyLeuLeuSerArgHisCysSerValAsn 2775
DB 7225 TATTCCTCATGAGGAGATACATCTTCTGTGACCTTCAGTTAGACAGTGCACAGCAAT 7284
QY 2776 GlyThrTrpThrGlySerAspProGlyCysLeuValIleAsnGlyAspProGlyTyrIle 2795
DB 7285 GGAATATGTCGAACTTACCTTACTTACTTAATCAATATCAAGTGTGGAGACCAAGTATA 7344
QY 2796 ProAlaAsnGlyLeuArgLeuGlyAsnAspPheArgTyrAsnLysThrValThrTyrGln 2815
DB 7345 CCAGCCATGAGCTGAGATGATGAGATGATTAATGTGTGACAAATATGTTCTTACATG 7404
QY 2816 CysValProGlyTyrMetMetGlySerHisArgValSerValLeuSerCysGlyThrLysAsp 2835
DB 7405 TGCCAGCGAGGCTACAGATGAATGAATGAGTCCAGAACTCAGAGCTTTGATCAATTAAT 7464
QY 2836 ArgThrTrpAsnGlyThrLysProValCysValAlaLeuMetCysLysProProProLeu 2855
DB 7465 GGCAATAGAGTGAATGAATGCCAACTGTAGACTTTTACCTGCCAATCTCTCCCAAG 7524
QY 2856 IleProAsnGlyLysValValGlySerAspPheMetTrpGlySerSerValThrTyrAla 2875
DB 7525 ATCTCTATGAGAGCTGAGGAAGAACAAATTTGCACTGGGCTTATGATATTAAGTACATC 7584
QY 2876 CysLeuGlyGlyTyrGlnLeuSerLeuProAlaValPheThrCysGlyGlyAsnGlySer 2895
DB 7585 TGTTCCTCAGAGCTATGAGCTATCTTCTCGCTGTGTTTGAACCTGTGAGGAATGTGAC 7644
QY 2896 TrpThrGlyLeuLeuProGlnCysPheProValPheCysGlyLysAspProGlyValProSer 2915
DB 7645 TGAGTGTGTGAAGTACCGAGTGTCTTAACAAATTTTGTGTGATACCTGTGATACCTGCC 7704
QY 2916 ArgGlyArgArgGlyLeuAspArgGlyPheSerTyrArgSerSerValSerPheSerCysHis 2935

```

Db 7705 CAAGAAAAAGAGAGCAAGGCTTATATATACAGTCAGAGCTTTCATTCAGCTGCAAT 7764

Qy 2936 ProProLeuValLeuValGlySerProArgArgPheCysGlnSerAspGlyThrThrSer 2955

Db 7765 TTTCCTTCATTAATTAAGGATCAACACCAAGATATGTCAAGACAGATGACATTTGAGT 7824

Qy 2956 GlyThrGlnProSerCysGlyLeuAspProThrLeuThrThrCysAlaAspProGlyValPro 2975

Db 7825 GGTTCATCACCTCACTGATAGAGCTTACCAAACTCTTGTGAAACCCAGAGTGCTGT 7884

Qy 2976 GlnPheGlyIleGlnAsnAsnSerGlnGlyThrGlnValGlySerThrValLeuPheArg 2995

Db 7885 CGGATGATCTCGAACAATACATTCGATTCAGATGAGCAAGTGTGTACATTCAT 7944

Qy 2996 CysGlnIleValGlyThrLeuLeuGlnGlySerThrThrArgThrCysLeuProAsnLeuThr 3015

Db 7945 TGCAAAAGAGACACCTTCTCCAAAGGCTTACAAACAGCACCTCTCTCCATGATTTACG 8004

Qy 3016 TrpSerGlyThrProProAspCysValProHisCysArgGlnProGlyThrProThr 3035

Db 8005 TGGAGTGGGATTCAGCTGATGATGATCCCAAGCTGTAAACAGCCAGAACTCTGCT 8064

Qy 3036 HisAlaAsnValGlyAlaLeuAspLeuProSerMetGlyThrThrLeuIle 3052

Db 8065 CATCAATATGTCGTAGGATGAGCTTCCATCTCATGGGTATACATGATT 8115

RESULT 14

AX537467

LOCUS AX537467 6004 bp DNA linear PAT 23-NOV-2002

DEFINITION Sequence 38 from Patent WO2070709.

AX537467

VERSION AX537467.1 GI:25269270

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

1 Lal, P.G., Baughn, M.R., Yao, M.G., Watia, N.K., Elliot, V.S., Xu, Y., Honchell, C.D., Yue, H., Ding, L., Gietzen, K.J., Ison, C.H., Lu, D.A., Hefalla, A.J., Gandhi, A.R., Thangavelu, K., Sanjanvala, M.M., Tang, Y.T., Ramkumar, J., Griffin, J.A., Swarnaker, A., Azimzal, Y., Sapperstein, S.K., Burford, N., Lee, E.A., Lu, Y., Tran, U.K. and Marquis, J.P.

TITLE Molecules for disease detection and treatment

JOURNAL Patent: WO 02070709-A 38 12-SEP-2002;

FEATURES

Source Location/Qualifiers

1..6004

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="Incyle ID No: 55061615CB1"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 6004

Score: 9077.50 Matches: 1720

Percent Similarity: 86.47% Conservative: 5

Best Local Similarity: 86.22% Mismatches: 8

Query Match: 53.44% Indels: 265

DB: 6 Gaps: 8

US-10-016-248-2 (1-3104) x AX537467 (1-6004)

Qy 859 AsnSerValThrGlyThrGlnGlyThrLeuLeuSerProAsnProValAsnTyAsn 878

Db 75 AATTCACATCAAGACCTCA---GGTACTTG---CTGTCCCAACGTC---CCCGTAATACAT 129

Qy 879 AsnAsnHisGlyCysIleTySerIleGlnThrGln-Pro-GlyValGlyIleGlnLeu 898

Db 130 GACAAATCTGAATGAT-TACTCCATGACAGCCAGTCCAGGAGTAACGGAATTCAGCTGA 168

Qy 898 YsAlaArgAlaPheGlnLeuSerGlnGlyAspValLeuIleValTyAspGlyAsnAsn 918

Db 189 AAGCCAGGCAATTCGAATCTCCGAGAGATGCTCCAAAGTTATGATGCAACAACA 248

Qy 918 snSerAlaArgLeuLeuGlyValPheSerHisSerGlnMetMetGlyValThrLeuAsn 938

Db 249 ACTCCGCCCGCTTGGCTGGAGGTTTTTTCATCTTCAGATTCAGATGAGGGGCTTTGAACA 308

Qy 938 erThrSerSerSerLeuThrLeuAsnPheIleThrAspAlaGlnAsnThrSerIleGlyP 958

Db 309 GCACATGCACACAGCTGTGGCTGTGATTCATCTGATGCTGAAACACAGCAAGGCT 368

Qy 958 heGlnLeuHisPheSerSerPheGlnLeuLeuIleValCysGlyAspProGlyThrProIle 978

Db 369 TTGAACGCACTTTTCCAGCTTTGAATCTCAATATGAGACCCAGAAACCCCAACT 428

Qy 978 heGlyTyThrValHisAspGlnGlyHisPheAlaGlySerSerValSerPheSerCysA 998

Db 429 TTGGCTCAAGAGTTCATGATGAGAGTCATTTTGGAGGAGACTCCGTCTTCAGCTGTG 488

Qy 998 sProGlyTyThrSerLeuAspGlySerGlnGlnLeuLeuCysLeuSerGlyGlnArgArg 1018

Db 489 ACCCTGGATACAGCTGCGGGGTAGTGAGAGCTGTGTCTGATGAGAGAGCGCGGA 548

Qy 1018 hTrpAspArgProLeuProThrCysValAlaGlnCysGlyGlyThrValArgGlyGln 1038

Db 549 CTGGGACCGGCTCTCTCCCACTGTCTCCCGAGTGTGAGGACAGTGAAGAGAGG 608

Qy 1038 aISerGlyGlnValLeuSerProGlyTyThrProAlaProTyArgIleHisAsnLeuAsnCysI 1058

Db 609 TGTGGGGGAGGAGTGTGTACACCGGGGTATCAGCTCCCTTGAACAACAATCTCACTGCA 668

Qy 1058 leTrpThrIleGlnAlaGlnAlaGlyCysThrIleGlyLeuHisPheLeuValPheAsp 1078

Db 669 TCTGGACCATCGAAGCAGAGCGCGCTGACCAATGGGCTACACTTCTGTGTGTACA 728

Qy 1078 hArgGlnValHisAspValLeuArgIleTrpAspGlyProValGlnSerGlyValLeu 1098

Db 729 CAGAGAGGTTCCAGAGCTCTGCGATCTGGAGTGGGCTGTGGAGAGCGGGTTTTC 788

Qy 1098 euIleGlnLeuSerGlyProAlaLeuProIleAspLeuHisSerThrPheAsnSerValY 1118

Db 789 TGAAGGAGCTGAGTGGCGCGCTGCTGCCAAGAGCTGATACACTTCAACTCGGTG 848

Qy 1118 alleuGlnPheSerThrAspPheThrSerIleGlnGlyPheAlaIleGlnPheSer 1138

Db 849 TCCTGCAGTTCACACACTGACTTCTTACCCAGCAAGCGGCTTGGCATTCATTTTCAG 908

Qy 1138 aISerThrAlaThrSerCysAsnAspProGlyIleProGlnAsnGlySerArgSerGlyA 1158

Db 909 TGTCCACAGCAGAGCTCTGCAATGACCTTGGATCCCGCAAAATGGAGTGGAGTGGTG 968

Qy 1158 sPseTrpGlnAlaGlyAspSerThrValPheGlnCysAspProGlyTyThrAlaLeuGln 1178

Db 969 ACAGTGGAGAGCGGGGACTCCACAGTGTCCAGTGTACCTGGGTACGGCTGACG 1028

Qy 1178 lYserAlaGlnIleSerCysValIleGlnLeuAsnArgPhePheThrGlnProSerPro 1198

Db 1029 GAAGTGAGAGATCAGCTGTGTGAGATCCAGAAACAGTGTCTTCTGGCACCCAGCGCG 1088

Qy 1198 roThrCysIleAlaProCysGlyGlyAspLeuThrGlyProSerGlyValIleLeuSerP 1218

Db 1089 CAACATGATGCTCTCTGCGGGGAGACTGACAGGACCATCTGGAGTATCTCTCAC 1148

Qy 1218 roAsnTyProGlnProTyProProGlyIleGlnCysAspTrpIleValThrValSerP 1238

Db 1149 CAATTAACCAAGAACCTTACCGCCAGGCAAGAGTGTACTGGAAGTACACGTCTCAC 1208

Qy 1238 roAspTyThrValIleAlaLeuValPheAsnIlePheAsnLeuGlnProGlyTyAspPhe 1258

Db 1209 CAGACTACAGCATCGGCTGTGTATTTAACTTTAACTGGAGGCTGTGATGACTTCC 1268

Qy 1258 euHisIleTyAspGlyArgAspSerLeuSerProLeuIleGlySerPheTyGlySerG 1278


```

Db      1269 TCATATCTACGACGCGAGCTCTCTAGGCCCTCTCATAGAAAGCTTCTATGGCTCC 1328
Qy      1278 InleuProglYarqIIeGIuSerSerSernSerleuPheleuA1aPheArSeraPA 1238
Db      1329 AGCTCCAGCGCGCATTTGAAGACAGACAGACAGCTCTCTCCGCTTCCGACGAGT 1388
Qy      1298 lAserValSerAmaIagIyPheValIleAspTYrThGluAsnProArqGluSerCySP 1318
Db      1389 CAtCTGTGACATGCTGGCTTCTGCTATTTGACTATACAGAAAACCGCGGAGATCATGTT 1448
Qy      1318 heAspProGIySerIleYsaAngIYrThArqValGIySerAspLeuYleuGIySers 1338
Db      1449 TTGATCTGGTTCCATCAAGACGCGCACACGGGTGGGGTCCGACCTGAAGCTGGGCTCT 1508
Qy      1338 eYValThTYrTYrCYseHIsagIyYrGIuValGIuGIyThSerThrIeuserCyS 1358
Db      1509 CCGTCACTACTACTACGCGACCGGGGGCTACGAAGTTGAGGGGCACTCGACCTGAGCTGCA 1568
Qy      1358 lEleuGIYProAspGIyYsPProValTPAsmaAsnProArqProValCYsThrAlaProc 1378
Db      1569 TCCTGGGGCTGTGATGGAAAGCCGTGTGMAACATCCCCGCCAGTCTGACACAGCCCCCT 1628
Qy      1378 YsGIyGIYGIuTYrValGIySerAspGIyValIleuSerProAsnTYrProGIuAsnT 1398
Db      1629 GTGGGGGACAGATATGTGGGTTCCGACGAGGTGCTTGTCCCCCACTACCCCCAGAACT 1688
Qy      1398 YrThSerGIYGIuIleCYseuTYrPheValThrValProLYsaAspTYrValIaPheG 1418
Db      1689 ACACAGTGCACAGATCTGCTTGTATTTTGTACTGTGCCCCAGAGACTATGTGTGTTG 1748
Qy      1418 lYGIuPheAlaPhePheHIsThrAlaLeuAsnAspValIaGIuValHIsAspGIYHIs 1438
Db      1749 GCCAGTTCTGCTTTTCAACAGCGCCCTCAACGACGTGTGAGAGTTCAACAGCGCCACA 1808
Qy      1438 eRGIuHIsSerArGIuLeuSerSerIeuserGIYSerHIsThr----- 1452
Db      1809 GCCAGCACTCGGGGCTCTCAAGCTCCCTCTGGGGCTCCATACAGATATCCGGGGCTCG 1868
Qy      1452 ----- 1452
Db      1869 CCACTGTGGGATGTGTGGGCGGGGGCATACGTCGGCTAAAGAAAGAGGCTCTTA 1928
Qy      1452 ----- 1452
Db      1929 GAAGCACCCATGCGCGAGGTGGAACTTACGCTCTGCGTCTGTGTTCTGTG 1988
Qy      1452 ----- 1452
Db      1989 CTGTGTCAACAGCAGACGACGCTGTGAGAGCTCCAACTAGCGGGGCTTTCAGACGT 2048
Qy      1452 ----- 1452
Db      2049 GCCCTACCCAGACTGTGTCTACACGCCCCCTTGTGTGTAGCCTTCTCTGTTGAATG 2108
Qy      1452 ----- 1452
Db      2109 GCAACTACACTATTGGCTGACAGGTCCAAGTTGTGTCTCTCCCTGGCCATCTGTA 2168
Qy      1452 ----- 1452
Db      2169 CTGCACCAAGACAGATATACCTTGTGTCTGTCTCAAAAAGCTGTACCTACCTGG 2228
Qy      1453 -----GlyGIuSerLeuProleuAlaThSerAnGIuValIleuIlel 1467
Db      2229 TTTCTGTGCGCATGACGAGATCACTGCTTGGCCACTCAATCAAGTTTCATTA 2288
Qy      1467 YsPheSerAlaYsGIYleuAlaProAlaArGIYpHeHIsPheValTYrGIu----- 1484
Db      2289 AGTTACGCCCAAGAGGCTCTGCACACGACAGAGGCTTTCATCAACCAAGGTATAG 2348
Qy      1485 -----AlaValProArqThSerAlaThrGIuCYseSerValProG 1499

```

```

Db      2349 AGGACATGACGCGCGAGCGTTCTCGAACCGACCGCACGACGTCTGTGCCGG 2408
Qy      1499 lUProArqTYrGIYLYsaTGIeUgIYSerAspPheSerValGIYAlaIleValArqPheG 1519
Db      2409 AACCCCGCTATGGCAAGAGGCTGGGAGTACTTCTGGTGGGGGCATCTGCCGTTGG 2468
Qy      1519 lUCYsaAsnSerGIYTYrAlaLeuGIuGIYSerProGIuIleGIuCYsLeuProValProG 1539
Db      2469 AATGCAACTCCGGCTATGCTCGACAGGGGTCCCGAAGATTCAGATGCCCTCTGTGCC 2528
Qy      1539 lValAlaLeuAlaGIuThPAsnValSerAlaProThCYsValIaIProCYsGIYIYAsnL 1559
Db      2529 GGAGCTTGGCCCAATGAATATGCTCAGCCGCCACAGTGTGTGTGCTGTGAGGCAACC 2588
Qy      1559 euThrGIuArqArqGIYThrIleLeuSerProGIYpHeProGIuProTYrIleuAsnSerL 1579
Db      2589 TCACAGAGCGAGGGGACCAATCTGTCCCTCGGCTTCCAGAGCCGTACTCAACAGCC 2648
Qy      1579 euBsnCYsValITrIYsIleValIaIProGIuGIYAlaGIYIleGIuIleGIuValIs 1599
Db      2649 TCAACTGTGTGTGAAGATCGGTGCCCGGAAGCGCTGGCATCCAGATCCAAATTGTCA 2708
Qy      1599 eYpHeValThrGIuGIuAsnThPAspSerIeUgIuValPheAspGIYAlaAspAsnThrv 1619
Db      2709 GTTTTGTGACAGACAGAACTGGGACTGTGGAAGTATTTGATGTGTCAGATACATG 2768
Qy      1619 aIThrMetIeUgIYSerPheSerGIYThrThrValProAlaLeuLeuAsnSerThSera 1639
Db      2769 TAACCATGTGGGAGTTTCTCAGAGAACACGTCGCCCTTCTGAACAGCACTTCA 2828
Qy      1639 sRGIuLeuTYrIleuHIsPheTYrSerAspIleSerValSerAlaAlaGIYpHeHIsleuG 1659
Db      2829 ACCAGCTTACCTTATTTACTAGATACACAGATATCGACGTGGCTTCCACTTGG 2888
Qy      1659 lUTYrLYsThrValGIYLeuSerSerCYsProGIuProAlaValProSerAnGIYValL 1679
Db      2889 AGTACAAACGCTGGGCTGAGAGTGTGTCCGGAACCTGCTGTGCCAGTACAGGGGTGA 2948
Qy      1679 YsThrGIYGIuArqTYrIleuValIAsnAspValIaIserPheGIuCYsGIuProGIYTYrA 1699
Db      2949 AGACTGGGAGCGCTACTGTGTGATGATGTGTGCTTTCCAGTGTGAGCGGGGATATG 3008
Qy      1699 lAleuGIuGIYHIsAlaHIsIleSerCYsMetProGIYThrThrValArqArqTrPAsnTYrP 1719
Db      3009 CCTTCAGGGGCGACGCCCACTCTCTGTGATGCCGGAACAGTGGCGGATGAACTAAC 3068
Qy      1719 RoPProleuCYsIleAlaGIuCYsGIYGIYThrValGIuGIuMetGIuGIYValIlel 1739
Db      3069 CTCTCACTCTGTATTTGACAGATGTGGGGAACAGTGAAGATGGAAGGGGTGATCC 3128
Qy      1739 euSerProGIYpHeProGIYAsnTYrProSerAnMetAspCYsETrIYsIleAlaL 1759
Db      3129 TGAAGCCCGGCTTCCAGGCACTACCCAGTAACTGACTCTCTCTGGAATAATAGCAC 3188
Qy      1759 euProValGIYpHeGIYAlaHIsIleGIuPheLeuAsnPheSerThrGIuProAsnHIsA 1779
Db      3189 TGCCGTGGGCTTGGAGCTCACATCAAGTCTCTGAATCTTCTCACAGGACCCAAACAG 3248
Qy      1779 sPTTYrIleGIuIleArqAsnGIYProTYrGIuThSerArqMetGIYArqPheSerG 1799
Db      3249 ACTACATGAAGAACTCGGAATGGCCCTATGAGACAGCGCATGATGGGAATTCAGTG 3308
Qy      1799 lYserGIuLeuProSerSerIeUleuSerThSerHIsGIuThrThrValTYrPheHIsS 1819
Db      3309 GAAGCAAGCTTCAAGCTCCCTCTTCCACGCTCCACAGACACACGAGTATTTCCACA 3368
Qy      1819 eYAspHIsSerGIuAsnArqProGIYpHeLYsleuGIuTYrGIu----- 1833
Db      3369 GCAGCACTCCAGATCGCGCAGATTCAGGTGAAGTATCAGATTGACTTACTCC 3428
Qy      1833 ----- 1833
Db      3429 ACCAGATTCTTCTCTGAGAGGTTTGTATCTCGAGGTTTGAAGAAACCAACTCA 3488

```

[illegible]

Oy	2176	helsvllleavgyserl a protyrcysseerleu prohdgla a proleuhisel ythei	2196
Db	4569	TCMAATCCGGCTATTCCA-----	4585
Oy	2196	leleuGLyglThSerThrGlnProGlySerIleHisphGlyCyaaAnaIaGLyT	2216
Db	4566	-----GGCGAACACAGACCCAGCCGGGGCTCATCTCACTTGGCTGCAACGCCGGCT	4640
Oy	2216	YrArgLeuValGlyHisSerMetAlaIleCystrHrArgHisProGlnGlyTrHisLeuT	2236
Db	4641	ACCGCTGTGGAGACACAGCATGGCCATCTGTACCCGGACACCCCAAGGCTACCACTGT	4700
Oy	2236	rpSerGluAlaIleProLeuGlySerGlnAlaLeuSerGlyGlyLeuProGluAlaProIySa	2256
Db	4701	GGAGGGAAGCATCCCTCTGTCAACCTCTTCTGTGGAGCTTCGAGGCCCCCAAGA	4760
Oy	2256	snGlyMetValPheGlyLyseGlyLurYrThrValGlyThrLyAlaValYrSerCyseSerg	2276
Db	4761	ATGGAATGTGTGGCGAAGAGTACACAGTGGGAACCAAGGCCATGTACAGCTGCAGTG	4820
Oy	2276	IuGlyYrHisAleuGlnAlaGLyAlaIaGLyAlaThrAlaGluCyAleuAspThrGlyLeuT	2296
Db	4821	AAGGTACACCTCCAGGCAAGCCCTGAGGCACTGCAGAGTGTGTGACACAGGCTAT	4880
Oy	2296	rpSerAsnArgAsnValProProGlnCyVal-----P 2307	
Db	4881	GGAGCAACCGCAATGTCCACCAAGTGTCTCTGAGCTCCGGGCATATGAGCGGGT	4940
Oy	2307	roValThrCyProAspValSerSerIleSerValGluHisGLyArgTrpArgLeuIleP	2327
Db	4941	CTGTGACTTGTCTGTATGTCAATGACATCAAGCGTGGAGACATGGCCGATGAGGCTTATCT	5000
Oy	2327	heGluThrGlnTrGlnPheGlnAlaGlnLeuMetLeuIleCyAspProGlyTYrTYrT	2347
Db	5001	TTGAGACACATATCAAGTCCAGGCCCAAGCGTCACTGTGACCTCGGCTACTACT	5060
Oy	2347	YrThrGlyGlnArgValIleArgCysGlnAlaAsnGlyLyETrpSerLeuGlyAspSerT	2367
Db	5061	ATACTGGCCAAAGGGTCACTCCGCTGTCAAGGCATATGGCAATGGAGCCTCGGGGACTCTA	5120
Oy	2367	hrProthCyArgIleIleSerGlyGlyLeuProIleProProAsnGlyHisArgI	2387
Db	5121	CGCCCACTCGCGAATCACTCTCTGTGAGAGCTCCCGATTCGCCCAATGGCCACCGCA	5180
Oy	2387	IeGlyThrLeuSerValTYrGlyAlaThrAlaIlePheSerCyAsnSerGlyTYrThrL	2407
Db	5181	TCGGAACACTGTCTGTCTACGGGGCAACAGCCATCTCTCTGCATTCGGGATACACAC	5240
Oy	2407	euValGlySerArgValArgGluCySerMetAlaAsnGlyLeuTrpSerGlySerGluValA	2427
Db	5241	TGTGTGGGCTCAGAGGTGTGTAGTGCATGGCCATGGGCTCTGGAGTGTCTGAACTCC	5300
Oy	2427	rgCybleu-----	2429
Db	5301	GCTGCTTGGCACTCAGACCAAGCTCCATCTTCTATTAAGCTCTCTGTGATGTAC	5360
Oy	2430	-----AlaGlyHisCyseGlyThrProGluProIleValAsnG	2442
Db	5361	TCTCTCCCACTCCCTCAACCAAGCTGACACCTGTGGGACTCTCTGAGCCCATTTGTCAAG	5420
Oy	2442	lyHisIleAsnGlyGluAsnTYrSerTYrArgIleSerValValTYrGlnCyAsnAlaG	2462
Db	5421	GACACATCAATGGGAGAACTACAGTACCGGGCAGTGTGTGTATCCATATGCAATGCTG	5480
Oy	2462	lyPheArgLeuIleGlyMetSerValArgIleCyseGlnGlnAspHisHisTrpSerGlyL	2482
Db	5481	GCTTCGCGTGTATGGGATGTGTGTGGCATCTGCACGACGAGATCACTGTGTGGGCA	5540
Oy	2482	YstrProPheCyVal-----	2487
Db	5541	AGACCCCTTTCTGTGTGATGATTTAAGACAGCATTTGCTGCTGCTGCTGCTGTGTGTG	5600
Oy	2488	-----ProIleThrCyseGlyHisProGlyAsnP 2497	

```

/evidence-not experimental
/product="ukf1A1894 protein"
/protein_id="BAC65849.1"
/db_xref="GI:28972866"
/translation="MARFFGDNFQJAGSSIVTICEEPIFKTQGETITTCILMDKVMWS
LAVPCGAPCGGHFAPASGVILSPMGQYKDSINCEVIAEAGHSIKTERPOTE
LNVDTLVEGPNLILSPILGSGYNGTVOPEQFSSNSITLLFTTNDNSRNNKGIHYE
STVNTVYGLDIPGIPVHGRRYGHDFSIGSTVYSSCDGPGVLSHEEPLCEKRWMSHP
LPTCAVAGSGDGRGSPGSGTLLSPGYPEPYNLNCWKTITPEAGAGTVOVSPFATHMNDSDPYDCG
HNYLITLMSGSPFOPIALRTSSELPTSTINAGLGNFRAQARTISDSISEVENTHDS
ENNERCECDPEIPIQTSRVFSPFVGDTLTFSSGLGRLEGSEIILCLDGGRRVMSAP
LERPCVAGSASATINNEGILLSPNPLVYENNEHCITSLQVADKQGINIARTPLHLAG
DLVLTQYDKDRTYHLGAFTGASMRGLTSTSTNGQMLSENSSECTDGPOLVYVFS
LHSHCEDPEIPQFQYKSIDQGHFAGSTIIGCNPGYTLHSSLSKCTGERRAMOVPL
PSCILAECCGRFKEGSGSRILSPGYPEPYNLNCWKTITPEAGAGTVOVSPFATHMNDSDPYDCG
HNYLITLMSGSPFOPIALRTSSELPTSTINAGLGNFRAQARTISDSISEVENTHDS
VALDGRGVPMNGTRDGRDREGRDTVYPCDGYELQOGERITTCIQVNRPIFWQSPS
PVICLAPCGGNLTGSSGRTLLSPNPPHYPIEBSRDQDMTIVNTDVISLARTISIEBNT
DPVLTIVDPSNLSPLISFODSKLPEIIESSNTMHARSDSIVSYTFHSEYAKL
RESGDENGINWGRILMDYKLTIVYVYCDAAVYLQGVSTLTCLMGDGRGRNARVL
PSCHAPGSRGTSGEETVLSPNPKYVSDNGVSIIVAKEEVAVRTSTGSSPSE
PRFRIRIGNDFAVGSLVEFCNPGYTLHSGRGLRGTGVNSLAKOMNDSILPTVICOGG
ILTRKRTGILSPGYPEPYNLNCWKTITPEAGAGTVOVSPFATHMNDSDPYDCG
DMNAPRLGYSYGTITLPLNSTSNLNIAPNOSDITVSAGAFHLBETATIGLDCPEPOT
PSSGILKIDGRVYMGDVYSPCCDGYSLQGHSHITTCMGRVKRMNRPITCLVACGGAM
SDSGVILSGPRENTPSSLDCTWTITKLPIGFEVHLQFVNFEFTETIHDLVYRGSSE
ISTVIGLSDGPQIPSSLPSTTHETSLYFHSIDSONKQGHIVQAYVOLQSCDPPFR
NGEIVIGNDPVGQTIISFCEPQYTLIGNSALTCLGSHNNMHPRLCGALCGNITAM
MGVITSPGYDEPYNFQDCEFMVLVPPNNGIYINFTVYQTEPIYDFTIVMGPONS
POLQCGNLTALESVYSTSNOLIKRHPSTSGPFLVSYHAYOLRVOCOPPVNAR
LITLDDFEIIGDILIRYOCLMGPVLGNALILTCRLGRLEPOMQDAPVCOVLCFANETL
DSTLSDPSPGYDPSPNLMOCAMSISVEGNVIMNFEVERPEKEDVQVYDGNITQ
SPVILISGDVYSAPVNTSNGHEVPLQMSADHNKNGKRRITVIAFYCSTSEPHGY
LIISQGGQLNSVTRACDNGFRVLGKSSVYTGSHMDAPVPAQALISCPKPA
PINGSILITDYLIVGRVTVFCNDGYRLSKELTTRTIGSHMDAPVPAQALISCPKPA
SINSFPLDHRGMRIVNGSHYEYKTKVFSFCDGVLGPASIECLPNTGMRTERP
COIISCELPPTPNKNGIKGTQTSYGSTALFTDCLDFMVLGASVRECLSGLMSGETR
CLAGHGIPELIVNQVILGENGYRDLVYVFCNIGMOPNRPILIGSSVTRICODHMSQGLP
CVAPSGHPSPIYGRGTSGNFPNDVYVFCNIGMOPNRPILIGSSVTRICODHMSQGLP
CYVNSCIDPGLIPANSKRREKIEHNGTCTQYVYFCNNGEYLPFSSVLTICOPNOMDK
PLCECIMIOGHPGIPRNVAVLSEKRTTSTVHISCTKRSILIGASHTCOLNHWGS
SOHSCGDDTTGTGDPGTGPHSROESDPRKTSVTRFACDITGVLIGSEBETCLNSGS
WTRORPCBAVOCNGPNTANKVPRIDGTTFSSTVILSCLEJGILSPVROCTANG
TWGSLPNCITLISCDPGLIPANGRLVGGDFFVUGVSNVMOCPGATIELNGSRVACTT
NGMWSGEVPCQLPKFCGDGPIPSQKRGKSTIYQSEVFSFNSCPBELLVGSSTPLCG
DGMWSSSPHCHIEPTRTSCENGPVRRHSONNTRFOVSGVYCHCKKGHILQGSTTR
TCLPDLTWSGIDPEICIPHSCPORESPRANVYGMDLBSHGTTLYTCOPNFFLAGTGE
HVRCSRDNMTKRVCEYCEAGSKILVQDRPALGTPSPULSVPDVVFPIATIMKGSYNF
KKRKQMTITVTSFNASTRVNATVLSNDBMLLSQGVKSOEARIMLHITVILIKPVA
SVAKMEENWMDGFSVASEPDAITVFGKIDGYKGDYVGLRGILMTSEBSNSNP
HGNSSSVAILVFPFALIFAGFGFYLYKORTAPKTYGTSCHENNNGQAPENEM
YDNTAKSVGKAVREDPMLNTVCTMV"

```


[illegible]

[illegible]

Dd	5372	TTGATCATCGAAGATGGAGATTTGGATTCACATTATGAAATCAAAACCTAAGGTG	5431
Qy	2338	MetLeuIleCyAspProGlyTyrTyrTyrThrGlyGlnArgValIleArgCyGlnAla	2357
Dd	5432	GTTTTCAGCTGTGATCCGGTTTATCATGAGACTTGGCCCTGCTCTATTGAATGCTTCCT	5491
Qy	2358	AsnGlyLysTrpSerLeuGlyAspSerThrProThrCysArgIleIleSerCyGlyGlu	2377
Dd	5492	AATGGTACTTGGATGGAGAACTGAAAGAACCGTACTGCCAAATCATCTCCTGGAGAG	5551
Qy	2378	LeuProIleProProIleAsnGlyHisArgIleGlyThrLeuSerValTyrGlyAlaThrAla	2397
Dd	5552	TTACCAACACTCCCAATGAGAAATTAAGATTGGAATCAAACTTCCTATGGCTCCAGCC	5611
Qy	2398	IlePheSerCyAsnSerGlyTyrThrThrLeuValGlySerArgValIArgGluCyMetAla	2417
Dd	5612	ATCTTCACATGTGATTTGGAGTTTCAATGCTTGTGGGCTCTGCTGTACCGGAATGCTCTCT	5671
Qy	2418	AsnGlyLeuTrpSerGlySerGlyValArgCysLeuAlaGlyHisCyGlyTyrProGlu	2437
Dd	5672	TCTGGTCTGTGATGGATGATCAGAAACCAAGATGTCTACCGGGTCACTGGAAATCCAGAA	5731
Qy	2438	ProIleValAsnGlyHisIleAsnGlyGluAsnTyrSerTyrArgGlySerValTyr	2457
Dd	5732	CTCATTTGGATGGGCAAGTCATTTGGAGAAATTAATGATATAGACACAGTTGTATAT	5791
Qy	2458	GlnCysAsnAlaGlyPheArgLeuIleGlyMetSerValArgIleCysGlnGlnAspHis	2477
Dd	5792	CAGTGTAACCTGGCTTCCTGCTCATATTTGGTTCTTCTGTAGAGATATATCAGCAGATCAC	5851
Qy	2478	HisTrpSerGlyLysTrpPhePheCysValProIleThrCysGlyHisProGlyAsnPro	2497
Dd	5852	AACGGTCTGGTCAAGCTCCCATCTGTGTGGCTGTATAGCTGTGGTCAACCTGGATGCCA	5911
Qy	2498	ValAsnGlyLeuThrGlnGlyAsnGlnPheAsnLeuAsnArgValValLysPheValCys	2517
Dd	5912	ATTATGAGAAACAAGATGGAAATGGGTTCAACTTATATGATGGTGAACATTTCTTCGC	5971
Qy	2518	AsnProGlyTyrMetAlaGluGlyAlaAlaArgSerGlnCysLeuAlaSerGlyGlnTrp	2537
Dd	5972	AATATTGGGATCTTATGACAAAGACCAACAAGACACAGTGTCAAGGCTTAACAGCAATGG	6031
Qy	2538	SerAspMetLeuProThrCysArgIleIleAsnCyThrAspProGlyHisGlnGluAsn	2557
Dd	6032	AGCATCTCTCCACAGTGTGCAAGGTGTGAATCTGTATCTGTGAATCCAGCCAAAT	6091
Qy	2558	SerValArgGln-----ValHisAlaSerGlyProHisArgPheSerPheGlyThr	2574
Dd	6092	TCCAAAGAGAAACCAAAATAGAACATGAAAT-----TTACCTTAATGGCACT	6139
Qy	2575	ThrValSerTyrArgCysAsnHisGlyPheTyrLeuGlyThrProValIleSerCys	2594
Dd	6140	GTTGGATTTCTTAAGCTGCACCCAGGCTATTTTTTATTTGGCTCTTCAGTTTGAATATGT	6199
Qy	2595	GlnGlyAspGlyThrTrpAspArgProArgProGlnCysLeuLeuValSerCyGlyHis	2614
Dd	6200	CAGGCAAAATGACAAATGGGCAAAACCTTTACACGAATGATCATGATTGACTGTGGACAC	6259
Qy	2615	ProGlySerProProHisSerGlnMetSerGlyAspSerTyrThrValGlyAlaValAla	2634
Dd	6260	CTGTGGCAATTCCTCTTAATGCTGTCTGTCTGGAGAGAAATATCTTTTGGGTCAACAGTT	6319
Qy	2635	ArgTyrSerCysIleGlyLysArgThrLeuValGlyAsnSerThrArgMetCysGlyLeu	2654
Dd	6320	CATTATTCCTGCACAGAAAGCGTTCTCTTTAGCGCAGGACATCCAGAACTTGCAGTTA	6379
Qy	2655	AspGlyHisTrpThrArgLysLeuProHisCysSerGlyThrSerValGlyValCysGly	2674
Dd	6380	AATGAGACTTGGATGGATGTCAACCTCATTTGTCAGGTATATCTATGTTACTATGTGGG	6439
Qy	2675	AspProGlyLysProAlaHisGlyLysArgLeuGlyAspSerPheAspProGlyThrVal	2694

Db 6440 GATCAGGTACACCTGTCATGGCTCCAGACAGAAAGTATTTCAGAACTAAAGCACT 6499
Qy 2695 MetArgPheSerCysGluAlaGlyHisValIleuAArgGlySerSerGluArgThrCysGln 2714
Db 6500 GTACGTTTGGCTGCTGATACCTGGATACATCTTTATGGCTCAGAAAGAACATGTTTA 6559
Qy 2715 AlaAsnGlySerTyrSerGlySerGlnProGluCysGlyValIleSerCysGlyAsnPro 2734
Db 6560 TCCAAATGGCAGTTGACCTGGAGACAGCCAGAGTCCAAGCTGTGCATGTGTATCCCA 6619
Qy 2735 GlyThrProSerAsnAlaArgValIlePheSerAspGlyLeuValPheSerSerIle 2754
Db 6620 GGAACCAACAGCCAAATGGGAAGCTTTCGAATTGATGGTACCAATTTCCAGTTCTGTC 6679
Qy 2755 ValTyrGluCysArgGluGlyTyrTyrAlaThrGlyLeuLeuSerArgHisCysSerVal 2774
Db 6680 ATCTATTTCTGCTGGAAAGCTATATCTTTTCAGSACCTCTGTTAGACAGTGACCTGCC 6739
Qy 2775 AsnGlyThrTyrThrGlySerAspProGluCysLeuValIleAsnCysGlyAspProGly 2794
Db 6740 AACGGAAACATGGCTGGAAGTCTGCCCAACTGTACATCATCAGTTGTGGAGACCAAGT 6799
Qy 2795 IleProAlaAsnGlyLeuArgLeuGlyAsnAspPheArgTyrAsnLysThrValThrTyr 2814
Db 6800 ATACCAAGCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6859
Qy 2815 GlnCysValProGlyTyrMetMetGluSerHisArgValSerValLeuSerCysThrLys 2834
Db 6860 ATGTGCCAGCCAGGCTATACCATAGTTAAATGATTAATGATCCAGATTCAGAACATGTACACT 6919
Qy 2835 AspArgThrTyrPheAsnGlyThrLysProValCysLysAlaLeuMetCysLysProProPro 2854
Db 6920 AATGGACATGAGTGAAGTGCATGCCAACATGATGAGCTGTACCTGCTCACTCCCTCC 6979
Qy 2855 LeuIleProAsnGlyLysValIleGlySerAspPheMetTyrGlySerSerValThrTyr 2874
Db 6980 CAGATCTCTATGGAAGCTGGAGAGAACAAATTTCACTGGGCTTTAGATATTAGCTAC 7039
Qy 2875 AlaCysLeuGluGlyTyrGlyLeuSerLeuProAlaValPheThrCysGluGlyAsnGly 2894
Db 7040 ATCTGTTCTGCAGGATATGAGCTGCTCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7099
Qy 2895 SerTyrThrGlyGluLeuProGlnCysPheProValPheCysGlyAspProGlyValPro 2914
Db 7100 ACCTGGAGTGAAGGTCCGCGAGTCTTACCAAAAGTTTGTGTGATCCTGTGATATACCT 7159
Qy 2915 SerArgLysArgArgLysAspArgGlyPheSerTyrArgSerSerValSerPheSerCys 2934
Db 7160 TCCCAAGAAACGAGAAAGCAAAAGCTTATATATCATGAGCTGAGTTTCACTTCACTGC 7219
Qy 2935 HisProProLeuValLeuValGlySerProArgArgPheCysGlnSerAspGlyThrTyr 2954
Db 7220 AATTTCTCCCTTCACTAGTGGATCAAGCAGAGGTTATGCCAAACTGATGGCACTTG 7279
Qy 2955 SerGlyTyrGlnProSerCysIleAspProThrLeuThrCysAlaAspProGlyVal 2974
Db 7280 AGTGGCTCATCACTCACTGATGAACCTTACAGAACGTCCTGTGAAAACCTGGAGTGA 7339
Qy 2975 ProGlnPheGlyIleGlnAsnAsnSerGlnGlyTyrGlnValIleGlySerThrValLeuPhe 2994
Db 7340 CCAAGGATGAGTCAAAACAAACACATTTGATTTCAAGTAGAAGTGTGTGGCAGTTTC 7399
Qy 2995 ArgCysGlnLysGlyTyrLeuLeuGlnGlySerThrThrArgThrCysLeuProAsnLeu 3014
Db 7400 CACTGCAAAAGGACACCTGCTTCAAGGATCTACAAACAGAACTTGCTCTGACCTG 7459
Qy 3015 ThrTyrSerGlyThrProProAspCysValProHisHisCysArgGlnProGluThrPro 3034
Db 7460 ACATGGAGTGAATTAACATGAGTCACTCACTCACTCACTCACTCACTCACTCACTCACT 7519
Qy 3035 ThrHisAlaAsnValGlyAlaLeuAspLeuProSerMetGlyTyrThrLeuIle 3052
Db 7520 GCTCATGCAAAAGTATGAGAAATGACCTTCCCTCCCACTGGGTACACACTCATC 7573

Search completed: October 18, 2004, 23:47:10
Job time : 28524 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 10:24:56 : Search time 95 Seconds
(without alignments)
10564.019 Million cell updates/sec

Title: US-10-016-248-2

Perfect score: 16985

Sequence: 1 MAGAPPALLPCLISDCC.....RSGPVDPSTLPGRHSPK 3104

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1360919 seqs, 32331874 residues

Total number of hits satisfying chosen parameters: 1360919

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : Listing first 45 summaries

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

19: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16985	100.0	3104	15	US-10-016-248-2 Sequence 2, Appl1
2	14142.5	83.3	2669	15	US-10-016-248-4 Sequence 4, Appl1
3	11305	66.6	3564	15	US-10-016-248-45 Sequence 45, Appl1
4	10975	64.6	3508	15	US-10-016-248-46 Sequence 46, Appl1
5	10475.5	61.7	3389	15	US-10-016-248-47 Sequence 47, Appl1
6	10475.5	61.7	3389	16	US-10-408-765A-2286 Sequence 2286, Ap
7	9025.5	53.1	1958	15	US-10-467-433-18 Sequence 18, Appl1
8	7662.5	45.1	2352	14	US-10-276-934-13 Sequence 13, Appl1
9	7643.5	45.0	2305	14	US-10-276-934-14 Sequence 14, Appl1
10	7568.5	44.6	2008	14	US-10-276-934-11 Sequence 11, Appl1
11	7098	41.8	1826	14	US-10-276-934-9 Sequence 9, Appl1
12	6949	40.9	1800	14	US-10-276-934-10 Sequence 10, Appl1
13	6936	40.8	1783	14	US-10-276-934-12 Sequence 12, Appl1
14	4226	24.9	1048	15	US-10-016-248-49 Sequence 49, Appl1

15	4209	24.8	1043	15	US-10-016-248-48 Sequence 48, Appl1
16	4143	24.4	1274	15	US-10-467-042-11 Sequence 11, Appl1
17	3174	18.7	613	9	US-09-799-514-9 Sequence 9, Appl1
18	2639	15.5	529	16	US-10-398-037-2 Sequence 2, Appl1
19	2555.5	15.0	839	16	US-10-398-455-11 Sequence 11, Appl1
20	1864.5	11.0	3567	14	US-10-028-248A-47 Sequence 47, Appl1
21	1864.5	11.0	3567	15	US-10-107-782-47 Sequence 47, Appl1
22	1850.5	10.9	3594	9	US-09-911-842-4 Sequence 4, Appl1
23	1850.5	10.9	3594	13	US-10-150-821-4 Sequence 4, Appl1
24	1794	10.6	3571	9	US-09-911-842-2 Sequence 2, Appl1
25	1794	10.6	3571	13	US-10-150-821-2 Sequence 2, Appl1
26	1793	10.6	3570	14	US-10-028-248A-6 Sequence 6, Appl1
27	1793	10.6	3570	15	US-10-107-782-6 Sequence 6, Appl1
28	1790	10.5	3571	16	US-10-603-283-2 Sequence 2, Appl1
29	1789	10.5	3568	14	US-10-028-248A-8 Sequence 8, Appl1
30	1789	10.5	3568	15	US-10-107-782-8 Sequence 8, Appl1
31	1782	10.5	3557	14	US-10-295-027-430 Sequence 430, App
32	1782	10.5	3557	14	US-10-295-027-1297 Sequence 1297, Ap
33	1629	9.6	326	15	US-10-451-010-10 Sequence 10, Appl1
34	1555	9.2	2489	9	US-09-911-842-5 Sequence 5, Appl1
35	1555	9.2	2489	13	US-10-150-821-5 Sequence 5, Appl1
36	1470.5	8.7	3623	16	US-10-741-601-335 Sequence 335, App
37	1419.5	8.4	2050	9	US-09-898-570-22 Sequence 22, Appl1
38	1419.5	8.4	2050	10	US-09-839-446-22 Sequence 22, Appl1
39	1370	8.1	1251	15	US-10-311-623-12 Sequence 12, Appl1
40	1362	8.0	2044	15	US-10-276-774-2152 Sequence 2152, Ap
41	1353	8.0	1004	14	US-10-114-270-32 Sequence 32, Appl1
42	1350.5	8.0	2887	17	US-10-479-875-8 Sequence 8, Appl1
43	1306	7.7	505	15	US-10-108-260A-3257 Sequence 3257, Ap
44	1281.5	7.5	351	15	US-10-416-314-1 Sequence 1, Appl1
45	1266.5	7.5	1139	9	US-09-764-893-102 Sequence 102, App

ALIGNMENTS

RESULT 1
US-10-016-248-2
Sequence 2, Application US/10016248
Publication No. US20040033491A1
GENERAL INFORMATION:
APPLICANT: Alsebrook et al.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-218
CURRENT APPLICATION NUMBER: US/10/016, 248
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: 60/254, 329
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/291, 037
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/255, 648
PRIOR FILING DATE: 2000-12-14
PRIOR APPLICATION NUMBER: 60/297, 173
PRIOR FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: 60/309, 258
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/326, 393
PRIOR FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: 60/315, 639
NUMBER OF SEQ ID NOS: 167
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 3104
TYPE: PRT
ORGANISM: Homo sapiens
US-10-016-248-2
Query Match 100.0%; Score 16985; DB 15; Length 3104;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 MAGAPPALLPCLISDCCASQNRHVGVPSELVKKQLKSRGVKLPMSKDSQKTS 60

```
Db 1 MAGAPPALLIPESLISDCCASQNRHSAVGSBELVKQIELKSGVKLMPKMDSOITS 60
Qy 61 VLTQVSGQHNNCPDPGI PERGKRLSGPRLGSSVOFTCNBGYLDQSKRTTCKMVS DM 120
Db 61 VLTQVSGQHNNCPDPGI PERGKRLSGPRLGSSVOFTCNBGYLDQSKRTTCKMVS DM 120
Qy 121 PAAMSHRPVCRARMDAHLRGPISGITSPPNFIQYDNNACWMIITLAPSKVTKLAFE 180
Db 121 PAAMSHRPVCRARMDAHLRGPISGITSPPNFIQYDNNACWMIITLAPSKVTKLAFE 180
Qy 181 EPDLERGDTLTVGOGGODGQKTVLYMSQNASDSDPHTPGSRIPESNSGDIWROKTVL 240
Db 181 EPDLERGDTLTVGOGGODGQKTVLYMSQNASDSDPHTPGSRIPESNSGDIWROKTVL 240
Qy 241 EICROISSDARSVRSKPTSNANVELVAPGTEIEQSCDPCGIPAYGREGRFHHGD 300
Db 241 EICROISSDARSVRSKPTSNANVELVAPGTEIEQSCDPCGIPAYGREGRFHHGD 300
Qy 301 TLKFCQAPAFELVQKAITCQKNQMSAKKQGCVCSCFNFPSBGVULSPNYEDYGNH 360
Db 301 TLKFCQAPAFELVQKAITCQKNQMSAKKQGCVCSCFNFPSBGVULSPNYEDYGNH 360
Qy 361 LHCWVLIARBPESRIHLAFNDIDVBPQDFLVIKDGATAEAPVLGTSNGQLPSSITSSG 420
Db 361 LHCWVLIARBPESRIHLAFNDIDVBPQDFLVIKDGATAEAPVLGTSNGQLPSSITSSG 420
Qy 421 HVARLEFQTDHSTGRGFNITETTRHNECPDPVPVNGKRFSGSLQSSISFLCDEGF 480
Db 421 HVARLEFQTDHSTGRGFNITETTRHNECPDPVPVNGKRFSGSLQSSISFLCDEGF 480
Qy 481 LGTQSEETITCVLKEGSAVWMSAVLRCRAPCGHLTSPSGTILSPGMPGFKALASCAMV 540
Db 481 LGTQSEETITCVLKEGSAVWMSAVLRCRAPCGHLTSPSGTILSPGMPGFKALASCAMV 540
Qy 541 IEADPGYIKITTPRPFTEVNYDTLEVRDGTYSAPLIGVHGTVQDFLISITSNLYL 600
Db 541 IEADPGYIKITTPRPFTEVNYDTLEVRDGTYSAPLIGVHGTVQDFLISITSNLYL 600
Qy 601 FSTDKSHSDIGFOLRYETITLQSDHCLDPGIPVNGGRHNDPFYVAGALVTESSCSGTLSD 660
Db 601 FSTDKSHSDIGFOLRYETITLQSDHCLDPGIPVNGGRHNDPFYVAGALVTESSCSGTLSD 660
Qy 661 GEPLECEBNFQWSPALPSCCALCGGFTQSSSGTILSPGPDFFYNNLNCWTIETSHGK 720
Db 661 GEPLECEBNFQWSPALPSCCALCGGFTQSSSGTILSPGPDFFYNNLNCWTIETSHGK 720
Qy 721 VFETPFHTEHESGHDYLLITENSGFTQPLROLTGSRLPAPISAGLGNFPAQVRFTSDPS 780
Db 721 VFETPFHTEHESGHDYLLITENSGFTQPLROLTGSRLPAPISAGLGNFPAQVRFTSDPS 780
Qy 781 MSYEGFNITSEYDLBCEBEPVAVSIRKGLQFGVGDITLFCSPFGYRLGRTARTCG 840
Db 781 MSYEGFNITSEYDLBCEBEPVAVSIRKGLQFGVGDITLFCSPFGYRLGRTARTCG 840
Qy 841 GRRRLMSSPLPRCAEAGNSVTGTQGLTLPNPPVNNNHECTIYSIQTPQKGKQLKAR 900
Db 841 GRRRLMSSPLPRCAEAGNSVTGTQGLTLPNPPVNNNHECTIYSIQTPQKGKQLKAR 900
Qy 901 AFELSEGDVLKYVDGNNSARLLGVFSHSEMGMVTLNSTSSILMDFTIDANTSTGFL 960
Db 901 AFELSEGDVLKYVDGNNSARLLGVFSHSEMGMVTLNSTSSILMDFTIDANTSTGFL 960
Qy 961 HBSSEFELIKEDDQTPKFGYKVADEGHFAGSSVSFCSDPGYSLRGSEELLCISGERRTMD 1020
Db 961 HBSSEFELIKEDDQTPKFGYKVADEGHFAGSSVSFCSDPGYSLRGSEELLCISGERRTMD 1020
Qy 1021 RPLPTCAECCGVVREBSGVULSFGYAPAYEHNLNCITITAEAGCTTGLHFLVPTDE 1080
Db 1021 RPLPTCAECCGVVREBSGVULSFGYAPAYEHNLNCITITAEAGCTTGLHFLVPTDE 1080
Qy 1081 VHDVLRIMDGPVBSGVULKELSGPALPKDLHSTFNSVVLQFSTDFTSKQGFALQFVSST 1140
Db 1081 VHDVLRIMDGPVBSGVULKELSGPALPKDLHSTFNSVVLQFSTDFTSKQGFALQFVSST 1140

Db 1081 VHDVLRIMDGPVBSGVULKELSGPALPKDLHSTFNSVVLQFSTDFTSKQGFALQFVSST 1140
Qy 1141 ATSCNPDGIPONGSRSGDSEWADSTVFOCDPBYALQGSNAEISCVKLENRFQPSPPTC 1200
Db 1141 ATSCNPDGIPONGSRSGDSEWADSTVFOCDPBYALQGSNAEISCVKLENRFQPSPPTC 1200
Qy 1201 IAPCGDULTPSGVILSPNYPBEPYPPGKECDMWKTVSPYVVALVNIENLBERGYFLH 1260
Db 1201 IAPCGDULTPSGVILSPNYPBEPYPPGKECDMWKTVSPYVVALVNIENLBERGYFLH 1260
Qy 1261 YDGRDLSPLIGSFYSSQIPGRIESSNSLFLAFPSDASVNAAGVITYDTEPRESCFDP 1320
Db 1261 YDGRDLSPLIGSFYSSQIPGRIESSNSLFLAFPSDASVNAAGVITYDTEPRESCFDP 1320
Qy 1321 GSIKNGTIRVSGDLKSSVTTYCHGGEYEGSTSLCIGPBGKPPMNNPRVCTAPCGG 1380
Db 1321 GSIKNGTIRVSGDLKSSVTTYCHGGEYEGSTSLCIGPBGKPPMNNPRVCTAPCGG 1380
Qy 1381 QYVSGDGVVLSPNYPQNTYSGQICLYPVTVPKDYVVFQGFAPFTALNDVVEVHDGSHQ 1440
Db 1381 QYVSGDGVVLSPNYPQNTYSGQICLYPVTVPKDYVVFQGFAPFTALNDVVEVHDGSHQ 1440
Qy 1441 SRLSLSGSHGSESLPLATSNQVLIKPSAKLARGFHYQAVPRTSATQSSVBP 1500
Db 1441 SRLSLSGSHGSESLPLATSNQVLIKPSAKLARGFHYQAVPRTSATQSSVBP 1500
Qy 1501 RYGRKLGSDFSYGAIVRECNAGVALQGSPEIECLPVGALANQNVSAPTCVPCGMLT 1560
Db 1501 RYGRKLGSDFSYGAIVRECNAGVALQGSPEIECLPVGALANQNVSAPTCVPCGMLT 1560
Qy 1561 ERRGTLSPGPEPPLNSLNCYWKIVBPAGAGIQIQQVSPYTEQWDSLEVPDAGDNTVT 1620
Db 1561 ERRGTLSPGPEPPLNSLNCYWKIVBPAGAGIQIQQVSPYTEQWDSLEVPDAGDNTVT 1620
Qy 1621 MLGSEFGTTPALNSTNOYLHFYSDISVSAAGFHLXYTVGIGLSCPEBAVSNQYKT 1680
Db 1621 MLGSEFGTTPALNSTNOYLHFYSDISVSAAGFHLXYTVGIGLSCPEBAVSNQYKT 1680
Qy 1681 GERRYVNDVVSFOCEBPGYALQGAHISCMPTVRMNPBPLCIAOCGIVEMEGLVLS 1740
Db 1681 GERRYVNDVVSFOCEBPGYALQGAHISCMPTVRMNPBPLCIAOCGIVEMEGLVLS 1740
Qy 1741 PGFPNYPNMDGSKIALPYGPAHIQPLNFSTEPNHDIYEINNGPETSRRMGRSGS 1800
Db 1741 PGFPNYPNMDGSKIALPYGPAHIQPLNFSTEPNHDIYEINNGPETSRRMGRSGS 1800
Qy 1801 ELPSLSLSTSHETTVYFHSDSQNRPGFKLEYQAYELOECBPDPFPANGIVRGAGYVQ 1860
Db 1801 ELPSLSLSTSHETTVYFHSDSQNRPGFKLEYQAYELOECBPDPFPANGIVRGAGYVQ 1860
Qy 1861 SVTFECLPGYOLGHPVLTCOHGNRMWDHPLPKCEVPCGANTSSNGTYVSPGSPYS 1920
Db 1861 SVTFECLPGYOLGHPVLTCOHGNRMWDHPLPKCEVPCGANTSSNGTYVSPGSPYS 1920
Qy 1921 SSODCWLITVPIGHGRLNLSLQTEPSGDFITWGPQOATPRLGIVFRSMAKKTUVOS 1980
Db 1921 SSODCWLITVPIGHGRLNLSLQTEPSGDFITWGPQOATPRLGIVFRSMAKKTUVOS 1980
Qy 1981 SSNOVLLKFRHDAATGCIFAIAPSAYPLTKCPPTILPNAEVTYENDEFNIGDIVRYCL 2040
Db 1981 SSNOVLLKFRHDAATGCIFAIAPSAYPLTKCPPTILPNAEVTYENDEFNIGDIVRYCL 2040
Qy 2041 PGFTLVGNELITCKLGTGYLOPEBGPPICEVHCTNELTDSSTVILIOSYPSGYPFOTC 2100
Db 2041 PGFTLVGNELITCKLGTGYLOPEBGPPICEVHCTNELTDSSTVILIOSYPSGYPFOTC 2100
Qy 2101 SMIVRVPDYNISLTVYFPLSEKQYDEFELFDPSGQSPLKALSGNYSAPLVTSSNS 2160
Db 2101 SMIVRVPDYNISLTVYFPLSEKQYDEFELFDPSGQSPLKALSGNYSAPLVTSSNS 2160
Qy 2161 VYLKMSDHAYNKRGFRIRASAYCSLPAPLHGFLIGQTSYTOPGSIHFGCAGYRLVG 2220
Db 2161 VYLKMSDHAYNKRGFRIRASAYCSLPAPLHGFLIGQTSYTOPGSIHFGCAGYRLVG 2220
```

2221 HSMALICRHPGNYLWSEATPLCOALSCGLPEAPKNGVNGKEYVTCTKAIVSSEBYHL 2280
2221 HSMALICRHPGNYLWSEATPLCOALSCGLPEAPKNGVNGKEYVTCTKAIVSSEBYHL 2280
2281 QAGAEATAECLDGLTGLMNRNVPPOCVPTCPDVSSIVEHGRWMLIETQYQFOQMLI 2340
2281 QAGAEATAECLDGLTGLMNRNVPPOCVPTCPDVSSIVEHGRWMLIETQYQFOQMLI 2340
2341 CDPGYVYTGQVRVRCQANGKMSLCDSTPTCRIISGGLPIPPNGHRIIGTISVGAIAIFS 2400
2341 CDPGYVYTGQVRVRCQANGKMSLCDSTPTCRIISGGLPIPPNGHRIIGTISVGAIAIFS 2400
2401 CNGGYTVLGSVRVRCMANGMSSEVRCAGHCTPPIVNGHNGENYVRSVYVQCN 2460
2401 CNGGYTVLGSVRVRCMANGMSSEVRCAGHCTPPIVNGHNGENYVRSVYVQCN 2460
2461 AGFLLIGMSVRIQQDHHMSGKTPFCVPTICGHPGNVNGLTQGNQFNLDVYKVCNPG 2520
2461 AGFLLIGMSVRIQQDHHMSGKTPFCVPTICGHPGNVNGLTQGNQFNLDVYKVCNPG 2520
2521 YMAEGAARSOCLASGQMSDMLPTCRIINCTDPGHQENSVRQVHASGPHRESFGTTVSYRC 2580
2521 YMAEGAARSOCLASGQMSDMLPTCRIINCTDPGHQENSVRQVHASGPHRESFGTTVSYRC 2580
2581 NHGTYLLGTPLVSCQGGGTWDRPRPQCLLVSCGHPGSPPHSQMSGDSYTVGAVVRYSCIG 2640
2581 NHGTYLLGTPLVSCQGGGTWDRPRPQCLLVSCGHPGSPPHSQMSGDSYTVGAVVRYSCIG 2640
2641 KRTLVGNTMCGLDGHWMTGSLPHCSGTSVGVCGDPPIPAHGIILGDSPPGTVMRSCCE 2700
2641 KRTLVGNTMCGLDGHWMTGSLPHCSGTSVGVCGDPPIPAHGIILGDSPPGTVMRSCCE 2700
2701 AGHVLKSSERTCQANGSWSGSGQPECGVISCNPGTSPNARVYFSDGLVFSSIVYECRE 2760
2701 AGHVLKSSERTCQANGSWSGSGQPECGVISCNPGTSPNARVYFSDGLVFSSIVYECRE 2760
2761 GYVATGLLSRHCSVNGTMTGSDPECLVINCDBPGLPANGRLGNDPFTVNTVYQVCNPGY 2820
2761 GYVATGLLSRHCSVNGTMTGSDPECLVINCDBPGLPANGRLGNDPFTVNTVYQVCNPGY 2820
2821 MMESHRSVLSCTKDRITMNGTKPVCKALMCKPPLIPNGKVVGSDFPMGSSVTYACLEGY 2880
2821 MMESHRSVLSCTKDRITMNGTKPVCKALMCKPPLIPNGKVVGSDFPMGSSVTYACLEGY 2880
2881 QLSIPAVFTCEBNGSWTGBLPQCFPVFCGDPGVSRGRREDRGFSYSSVSFSCHPPLVL 2940
2881 QLSIPAVFTCEBNGSWTGBLPQCFPVFCGDPGVSRGRREDRGFSYSSVSFSCHPPLVL 2940
2941 VGSRRRCQSDGWTGTPQSPCIDPTLTTCADPGVPGQICQNSGQYGVGSTVLFRCKQGY 3000
2941 VGSRRRCQSDGWTGTPQSPCIDPTLTTCADPGVPGQICQNSGQYGVGSTVLFRCKQGY 3000
3001 LLOGSTTRTCLPNTLWMSGTPPDVPHHROPEPTPHANVGALDPPSMGYTLITPARASP 3060
3001 LLOGSTTRTCLPNTLWMSGTPPDVPHHROPEPTPHANVGALDPPSMGYTLITPARASP 3060
3061 SRVAPSTAPARMMAAGQASRPSANRSGPVGDPTLPGSHRSPKP 3104
3061 SRVAPSTAPARMMAAGQASRPSANRSGPVGDPTLPGSHRSPKP 3104

RESULT 2

US-10-016-248-4
; Sequence 4, Application US/10016248
; Publication No. US2004003491A1
; GENERAL INFORMATION:
; APPLICANT: Aliebrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-218
; CURRENT APPLICATION NUMBER: US/10/016, 248
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/254, 329

PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/291, 037
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/255, 648
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 60/297, 173
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/309, 258
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/326, 393
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/315, 639
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 157
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2669.
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-016-248-4

Query Match 83.3%; Score 14142.5; DB 15; Length 2669;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 2605; Conservative 1; Mismatches 2; Indels 59; Gaps 3;

1 MAGAPPALLPGLSLSDCCASNORHSVGVSELYVKQIETLSRGVKKMPKSDNSQKTS 60
1 MAGAPPALLPGLSLSDCCASNORHSVGVSELYVKQIETLSRGVKKMPKSDNSQKTS 60
61 VLTQVGVSGQNNMCPDGPPIPERGKRLGSDPRLGSSVQFTCNEGYDLOGSKRITCMKYS 120
61 VLTQVGVSGQNNMCPDGPPIPERGKRLGSDPRLGSSVQFTCNEGYDLOGSKRITCMKYS 120
61 VLTQVGVSGQNNMCPDGPPIPERGKRLGSDPRLGSSVQFTCNEGYDLOGSKRITCMKYS 120
121 FAAMSDHRPVCARMCDALHSGSGIITSNPIQYDNNAHQWIIITALNPSTYIKLAFE 180
121 FAAMSDHRPVCARMCDALHSGSGIITSNPIQYDNNAHQWIIITALNPSTYIKLAFE 180
181 EPLREGEYDPLTVYDGGQSDQDQKTVLYMSQNASDSHPHTGSRIPESMSGDIROKKTVL 240
181 EPLREGEYDPLTVYDGGQSDQDQKTVLYMSQNASDSHPHTGSRIPESMSGDIROKKTVL 240
181 EPLREGEYDPLTVYDGGQSDQDQKTVLYMSQNASDSHPHTGSRIPESMSGDIROKKTVL 240
241 EICRDISSDARSRSYKSPKTSNAVELVAPGTEIEQSGSCGDPGIPAYGRREGSRFHHGD 300
241 EICRDISSDARSRSYKSPKTSNAVELVAPGTEIEQSGSCGDPGIPAYGRREGSRFHHGD 300
241 EICRDISSDARSRSYKSPKTSNAVELVAPGTEIEQSGSCGDPGIPAYGRREGSRFHHGD 300
301 TLKFECPAPFELVQKALITCQKNNQMSAKKPGCVFSCFENFTSPSGVILSPNTPEDYGNH 360
301 TLKFECPAPFELVQKALITCQKNNQMSAKKPGCVFSCFENFTSPSGVILSPNTPEDYGNH 360
301 TLKFECPAPFELVQKALITCQKNNQMSAKKPGCVFSCFENFTSPSGVILSPNTPEDYGNH 360
361 LHCWVLLAPRERIHAFNDIVEPOFDLVTKDQATAPVILGTFSSGNQLSSITSSG 420
361 LHCWVLLAPRERIHAFNDIVEPOFDLVTKDQATAPVILGTFSSGNQLSSITSSG 420
361 LHCWVLLAPRERIHAFNDIVEPOFDLVTKDQATAPVILGTFSSGNQLSSITSSG 420
421 HVARLEFQDTHSGKSGFNITFTFRHNECPDGVPPVNGKRFSDSLQAGSSIFLCEDEGF 480
421 HVARLEFQDTHSGKSGFNITFTFRHNECPDGVPPVNGKRFSDSLQAGSSIFLCEDEGF 480
421 HVARLEFQDTHSGKSGFNITFTFRHNECPDGVPPVNGKRFSDSLQAGSSIFLCEDEGF 480
481 LGTQSEFTITCVLKEGSSVWNSAVLRCBACGHLTSPSGTILSPGMGFFYKALSCAWV 540
481 LGTQSEFTITCVLKEGSSVWNSAVLRCBACGHLTSPSGTILSPGMGFFYKALSCAWV 540
481 LGTQSEFTITCVLKEGSSVWNSAVLRCBACGHLTSPSGTILSPGMGFFYKALSCAWV 540
541 IEAQPGYPIKITDRFTEVNYDTLEVRDRTYSAPLIGYHGTQVPOFLISTSNLYYL 600
541 IEAQPGYPIKITDRFTEVNYDTLEVRDRTYSAPLIGYHGTQVPOFLISTSNLYYL 600
541 IEAQPGYPIKITDRFTEVNYDTLEVRDRTYSAPLIGYHGTQVPOFLISTSNLYYL 600
601 FSTDKSHSDIGPOLRYETITLQSDHCLDPGIPVNGQRHGNDFYVAGLVTFSCSGSYTLSD 660
601 FSTDKSHSDIGPOLRYETITLQSDHCLDPGIPVNGQRHGNDFYVAGLVTFSCSGSYTLSD 660
601 FSTDKSHSDIGPOLRYETITLQSDHCLDPGIPVNGQRHGNDFYVAGLVTFSCSGSYTLSD 660
661 GEPLECEPNFQWRALPSCFALCGGFTQSSGTTILSPGPDYFPNNINCTWIIETSHGKG 720
661 GEPLECEPNFQWRALPSCFALCGGFTQSSGTTILSPGPDYFPNNINCTWIIETSHGKG 720
661 GEPLECEPNFQWRALPSCFALCGGFTQSSGTTILSPGPDYFPNNINCTWIIETSHGKG 720

```

QY 721 VFPTHTFHLSESHDYLITENGSTFOTPLRLQTLGSRPLAPISAGLXGNFTTAQVRFISDPS 780
Db 721 VFPTHTFHLSESHDYLITENGSTFOTPLRLQTLGSRPLAPISAGLXGNFTTAQVRFISDPS 780
QY 781 MSYGEFINTSEIDLEPCEEPEVPAYSIRKGLQFVGDTLTFCSEFPYRLGRTARTICG 840
Db 781 MSYGEFINTSEIDLEPCEEPEVPAYSIRKGLQFVGDTLTFCSEFPYRLGRTARTICG 840
QY 841 GRRRLMSSPLPRCAVEGNSVTGOTLTLSPNFPVYNNNNHBCIYSIQTPGKGIQLKAR 900
Db 841 GRRRLMSSPLPRCAVEGNSVTGOTLTLSPNFPVYNNNNHBCIYSIQTPGKGIQLKAR 900
QY 901 AFELISGDVLKVDGNNNSARLLGVFSHSEMMGVTLNSTSSLMIDFTDAENTSKGPEL 960
Db 901 AFELISGDVLKVDGNNNSARLLGVFSHSEMMGVTLNSTSSLMIDFTDAENTSKGPEL 960
QY 961 HFSSEFELICEDGTGKFGYKXNDDEHFKGSSVSTSCDPGYSLRSEBELCLSGERRTWD 1020
Db 961 HFSSEFELICEDGTGKFGYKXNDDEHFKGSSVSTSCDPGYSLRSEBELCLSGERRTWD 1020
QY 1021 RPLPTCVACCGGTVRGEVSGOVLSPGYAPAYEHNNCTIWEABAGCTTGHFLVPTDEE 1080
Db 1021 RPLPTCVACCGGTVRGEVSGOVLSPGYAPAYEHNNCTIWEABAGCTTGHFLVPTDEE 1080
QY 1081 VHDVLRIMDGPVSESVTLKELSGPALPKDLHSTFNSVVLQFSTDFTSKQGFALQFVSST 1140
Db 1081 VHDVLRIMDGPVSESVTLKELSGPALPKDLHSTFNSVVLQFSTDFTSKQGFALQFVSST 1140
QY 1141 ATSCNDPGIPONGSSRSGDMEAGDSTVQCDDPGVALQGSARISCVKIENRFQOPSPPTC 1200
Db 1141 ATSCNDPGIPONGSSRSGDMEAGDSTVQCDDPGVALQGSARISCVKIENRFQOPSPPTC 1200
QY 1201 IAPCGGDLTGPGSVILSPNYPPEYPGKCECMKVTSPDVIALVFNIFLPEGYDLPHI 1260
Db 1201 IAPCGGDLTGPGSVILSPNYPPEYPGKCECMKVTSPDVIALVFNIFLPEGYDLPHI 1260
QY 1261 YDGRDSTLPLIGSFFGSQLPGRIBSSNSLFLARSDASVSNAGFVLDYENPRESCFDP 1320
Db 1261 YDGRDSTLPLIGSFFGSQLPGRIBSSNSLFLARSDASVSNAGFVLDYENPRESCFDP 1320
QY 1321 GSINKGTIVGSDLKLGSSVTVYCHQGYEVEGTSTLSCILGPDGKPVNNNRPVCTAPCGG 1380
Db 1321 GSINKGTIVGSDLKLGSSVTVYCHQGYEVEGTSTLSCILGPDGKPVNNNRPVCTAPCGG 1380
QY 1381 QYVSGDGVVLSBNYPONTSGOICLYPYTVPKDYVVGOPAFPHALINDVVEVHDGSHQ 1440
Db 1381 QYVSGDGVVLSBNYPONTSGOICLYPYTVPKDYVVGOPAFPHALINDVVEVHDGSHQ 1440
QY 1441 SRLLSLSGSHGTGSLPLATSNQVLIKFSAKGLAPARGFPVVOAVPRTSATQCCSVPEP 1500
Db 1441 SRLLSLSGSHGTGSLPLATSNQVLIKFSAKGLAPARGFPVVOAVPRTSATQCCSVPEP 1500
QY 1501 RYGRRLSDSFSVGLIYFECNSGYALQGSPEIECLPVPGALAOVNSAPTCVPCGANT 1560
Db 1501 RYGRRLSDSFSVGLIYFECNSGYALQGSPEIECLPVPGALAOVNSAPTCVPCGANT 1560
QY 1561 ERRGTILSPGPEPEYLSLNCWKIYVPEGAQIOIVVSPVTEOMNDSLEVPDQADNTY 1620
Db 1561 ERRGTILSPGPEPEYLSLNCWKIYVPEGAQIOIVVSPVTEOMNDSLEVPDQADNTY 1620
QY 1621 MLGSFSGTTPALINSTSNQVLIHFYSDISVSAGHLEKTVGLSCEPAVPSNGVKT 1680
Db 1621 MLGSFSGTTPALINSTSNQVLIHFYSDISVSAGHLEKTVGLSCEPAVPSNGVKT 1680
QY 1681 GERLVNDVVSFOCEGEYALQGHAAHISCMGTYRRANNYPPLICACCGTVEEMGVILS 1740
Db 1681 GERLVNDVVSFOCEGEYALQGHAAHISCMGTYRRANNYPPLICACCGTVEEMGVILS 1740
QY 1741 PGFPGVPSNMDSCWKIALPVGGAHIQELNSTEPNHDYIETIRNGPYETSMMKRFSGS 1800
Db 1741 PGFPGVPSNMDSCWKIALPVGGAHIQELNSTEPNHDYIETIRNGPYETSMMKRFSGS 1800
QY 1801 ELPSLSLSTSHETTVFHSQHSQNRPGFKLEYOAYELOECPPDEPPANGIVRAGAYNVQ 1860

```

```

Db 1801 ELPSLSLSTSHETTVFHSQHSQNRPGFKLEYOAYELOECPPDEPPANGIVRAGAYNVQ 1860
QY 1861 SVTEECPLRGYQLGHPVLTQHGHTNNMDHPLKCEVPCGANTSSNGTVYSGFPSYS 1920
Db 1861 SVTEECPLRGYQLGHPVLTQHGHTNNMDHPLKCEVPCGANTSSNGTVYSGFPSYS 1920
QY 1921 SSQDCWMLITVPIGHGVRNLNLSTQTEPSGDFITINDGQOOPAPRLGVTRTSMAKTUOS 1980
Db 1921 SSQDCWMLITVPIGHGVRNLNLSTQTEPSGDFITINDGQOOPAPRLGVTRTSMAKTUOS 1980
QY 1981 SSMOVLKTRDAATGCIPIALAFSAVPLTKCEPPTLLPNAEVTENEENFIDIVRYCL 2040
Db 1981 SSMOVLKTRDAATGCIPIALAFSAVPLTKCEPPTLLPNAEVTENEENFIDIVRYCL 2040
QY 2041 PGFTLVGNELTCKTGTYOFSGPPICGVHCPTEMLLTDSTGVILSOSYPSGYPQFOTC 2100
Db 2041 PGFTLVGNELTCKTGTYOFSGPPICGVHCPTEMLLTDSTGVILSOSYPSGYPQFOTC 2100
QY 2101 SWLVREPDYNI SLVVEYFLSEKQYDEFEIPDGPESQSPFLKALSGNYSAPLITVSSNS 2160
Db 2101 SWLVREPDYNI SLVVEYFLSEKQYDEFEIPDGPESQSPFLKALSGNYSAPLITVSSNS 2160
QY 2161 VYLKMSDHAIVNRKGFRTYSAFYCSLRAPLHGLIQTSTOPGSGIHFGCNAGYRLVG 2220
Db 2161 VYLKMSDHAIVNRKGFRTYSAFYCSLRAPLHGLIQTSTOPGSGIHFGCNAGYRLVG 2220
QY 2221 HSMALICRHPQGYHLMSEALPLCOALSGLPEAPKNGVWFGKBYVGTKAYVSCSEGYH 2280
Db 2221 HSMALICRHPQGYHLMSEALPLCOALSGLPEAPKNGVWFGKBYVGTKAYVSCSEGYH 2280
QY 2281 QMGAETAECLDTGLMSNRNVPQCV-----PVTCPDVSSI VEHGRMLIFETQY 2340
Db 2281 QMGAETAECLDTGLMSNRNVPQCVRESSGNGGGSVTCPEVSSI VEHGRMLIFETQY 2340
QY 2341 QFOAOLMLICDPGYYYTQQRVYRQOANGKWSIGDSTPCRIISGCELPBPNGHRTIGLS 2400
Db 2341 QFOAOLMLICDPGYYYTQQRVYRQOANGKWSIGDSTPCRIISGCELPBPNGHRTIGLS 2400
QY 2401 VYGATAIYSCNSGYTLVGSRYRECMANGLWSGSEVRCL----- 2460
Db 2401 VYGATAIYSCNSGYTLVGSRYRECMANGLWSGSEVRCLATQYTKLHSIFYKLLPVLSSPS 2460
QY 2461 LTKGHCTPBPYNGHNGENBYSYRSVYQCNAGFRLIGMSVRLCOODHMSGKTPFC 2520
Db 2461 LTKGHCTPBPYNGHNGENBYSYRSVYQCNAGFRLIGMSVRLCOODHMSGKTPFC 2520
QY 2521 VHVRCOQLLLLLLLCDDDDDEDDSGAITCGHGNPNVNGLTQGNQFRLANDVAVFCNPGY 2580
Db 2521 VHVRCOQLLLLLLLCDDDDDEDDSGAITCGHGNPNVNGLTQGNQFRLANDVAVFCNPGY 2580
QY 2581 MABGAANSQCLASQSWDMPTCRIINCTDPGHQENSVRQVHASGHRFSFGTTVSRCV 2640
Db 2581 MABGAANSQCLASQSWDMPTCRIINCTDPGHQENSVRQVHASGHRFSFGTTVSRCV 2640
QY 2641 HGFYLLGTPLYSCGSDGTMDRPPQCL 2608
Db 2641 HGFYLLGTPLYSCGSDGTMDRPPQCL 2667

```

```

RESULT 3
US-10-016-248-45
; Sequence 45, Application US/10016248
; Publication No. US20040033491A1
; GENERAL INFORMATION:
; APPLICANT: Alcobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-218
; CURRENT APPLICATION NUMBER: US/10/016,248
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/254,329
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/291,037

```

PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/255,648
PRIOR FILING DATE: 2000-12-14
PRIOR APPLICATION NUMBER: 60/297,173
PRIOR FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: 60/309,258
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/326,393
PRIOR FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: 60/315,639
PRIOR FILING DATE: 2001-08-29
NUMBER OF SEQ ID NOS: 167
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 45
LENGTH: 3564
TYPE: PRI
ORGANISM: Mus musculus
US-10-016-248-45

Query Match 66.6%; Score 11305; DB 15; Length 3564;
Best Local Similarity 63.8%; Pred. No. 0;
Matches 1960; Conservative 463; Mismatches 592; Indels 58; Gaps 8;

QY 1 MAGAPPAL- - - - - LFCSLISD- - - - - CCASNORSHVGVSELYKKQIELKRGV 47
DB 264 ISGTAEASIMLTGKNLPSVLISSKMWLRHFTSDSNHRRK-GRVAQFOVKKALIELKRGV 322
QY 48 KLPMSKNSOKTVLTVGVSGQHNMCPPGIGPERGKRLSDPFLSSVOFTNEGVDLQ 107
DB 323 KMLPSKSSHKNSVLTGCGVSLSDMCPDPGIPDNGRAGSDFVGVANVOFSCEDNYLQ 382
QY 108 GSKRITOMKYSDMRAWSDRPVRCARMCALHGRSGIITSPFPIYOYNNACWIT 167
DB 383 GAKGITOQRTVETLTAANNDRPICRATCGSNLKGPSGLTSPVVOYEDNACWIT 442
QY 168 ALNPSKYIKLAFEEFDELRGYDILTVGDGQDQDKTVLWMSQVACSDSPHTPSRIPE- 226
DB 443 TTDPDKIKLAFEEFELERGYDILTVGDAGKVDTRVLV- - - - - LGSSVPLD 492
QY 227 - - - - - SWSGDIWROKMTVLEICRDISSDARSQSVKSPKTSNAVELVARG- - - - - TELIQS 279
DB 493 IVSMNQMWLH- - - - - LQSDDS- - - - - IGSPFAVAVQELIKG 526
QY 280 CGDGIAYGRRESRPHHGDILKECOPAFELVYOKAITCQKNNMSAKKPGVSCF 339
DB 527 CGDGIAYGRGTSSFLHGDILTFECQAFELVGERVITCQKNNMSGKNSPSCVSCF 586
QY 340 NFTSPSGVLSNPYPEDYGNHLICVWLILARPESRIHLAFENDIDVEPQDFLVIKDGATA 399
DB 587 NFTAPEGIILSPNPEYGNMNCVWLILSPESRIHLIFDPRVEPQDFELAVKDGIS 646
QY 400 EAPVLGTFSGNQLPSSITSSGVARLEFQTDHSTGKRGFNITFTFPHNECPDPGVNG 459
DB 647 DITVLGTFSGNEVBAQALSSGHIIVLEFQSDHSTTGAGFNITFTFQNECHDGIPIVNG 706
QY 460 KRREDSQLQSSIFLDEGFLTGQSEETICVUKESVYMNKAVLACEAPCGHILSPS 519
DB 707 RRFEDRPLSSVSFHCDDGFKVQSSSITCILLQDANVWSSVTPPCCEAPCGHILSPS 766
QY 520 GTILSPMPGPKYKALSCAWIYEAQPGYIKITFDRKTEVNYNYTLEVRDQRTASALIG 579
DB 767 GVILPMPGPKYKALSCAWIYEAQPGYIKITFDRKTEVNYNYTLEVRDQRTASALIG 826
QY 580 VYHGTQVQPLISTSNLYLLFSTDKSHSDIGQLRYETITLQSDHCLDPGIPVNGRG 639
DB 827 EYHGTQVQPLISTSNLYLLFSTDKSHSDIGQLRYETITLQSDHCLDPGIPVNGRG 886
QY 640 NDFVGLVYTESCSGTYLSDGEPLCEPNFQWRALPSCALCGGFIQSSGTTILSPG 699
DB 887 SNFIRSTVTFSCDPGTYLSDDEPLVCKNQHMLAPSCALCGGFIHGSQTYLSPG 946
QY 700 PDFPNNUNCTWITETSHKGVFTFTPHLESCHDYLLITENGSTFOPLRQLTGSRIPA 759

DB 947 PDFPNNUNCTWITETSHKGVFTFTPHLESCHDYLLITENGSTFOPLRQLTGSRIPA 1006
QY 760 PISAGLYGNFTAAVQVITSDPSMSYEGENTFSSYDLEPCGEPPVAVSIRKQIOFGVDT 819
DB 1007 TIFAGLFGNFTAAVQVITSDPSMSYEGENTFSSYDLEPCGEPPVAVSIRKQIOFGVDT 1066
QY 820 LTFSCFPYVLEGTARITCIGRRRLMSPLPCVACGNSVYGTCTLLSPNFPVANN 879
DB 1067 LATTCQGYLBEATKTLTICGGRRLMSALPCVACGNSVYGTCTLLSPNFPVANN 1126
QY 880 NHECTYSIQTPGKIQKARAFELSEBDVLYKYGNNNSARILLGVPSSEMMGVTLNST 939
DB 1127 NHECTYSIQTPGKIQKARAFELSEBDVLYKYGNNNSARILLGVPSSEMMGVTLNST 1186
QY 940 SSSIMLDTITDAENTSKGFLHSSPELLKCEPDPGPKYKXKHDGHPFAGSSVSPGCDP 999
DB 1187 SSSIMLDTITDAENTSKGFLHSSPELLKCEPDPGPKYKXKHDGHPFAGSSVSPGCDP 1246
QY 1000 GYSLRGSEELLCSGERRTWDRPLPTCVACGCTVGEVSGVLYSPGPAPYEHNLNCIW 1059
DB 1247 GYSLRGSEELLCSGERRTWDRPLPTCVACGCTVGEVSGVLYSPGPAPYEHNLNCIW 1306
QY 1060 TIEAAGCTTGHPVLEEDVHDLRIMDPVYSGVLLKELSGPALPKDLSHPNSVL 1119
DB 1307 TIEAAGCTTGHPVLEEDVHDLRIMDPVYSGVLLKELSGPALPKDLSHPNSVL 1366
QY 1120 QESTDPTSKQFALIQSVSTATSCNDPGLPONGSSSGDWEAGDSIVPQCDGVALQGS 1179
DB 1367 QESTDPTSKQFALIQSVSTATSCNDPGLPONGSSSGDWEAGDSIVPQCDGVALQGS 1426
QY 1180 AELSCVKEIRPFMOQSPPTCIAPCGGLDLPSEGVLSIPVPPYPPGKCDMKVTVSPD 1239
DB 1427 AKITCQVQNNRFTWQDPPEPCIALCGGNTGPAVGLISPIYPPYPPGKCDMKVTVSPD 1486
QY 1240 VYIALVENIFNLEPGYDLHLYDGRDLSPLIGSPYQSLPGRIESSNSLFLAFRSDAS 1299
DB 1487 VYIALVENIFNLEPGYDLHLYDGRDLSPLIGSPYQSLPGRIESSNSLFLAFRSDAS 1546
QY 1300 VSNAGFYIDTNPRESCTPGRSITKNGTRGSLDKGSSVTVYCHGYEVGEGSTLSCL 1359
DB 1547 VSNAGFYIDTNPRESCTPGRSITKNGTRGSLDKGSSVTVYCHGYEVGEGSTLSCL 1606
QY 1360 GPDGKPYMNNPRVPCAPCGGQYVSGVLSIPVPPYPPGKCDMKVTVSPD 1419
DB 1607 GPDGKPYMNNPRVPCAPCGGQYVSGVLSIPVPPYPPGKCDMKVTVSPD 1666
QY 1420 FAFETHALNDVVEHDSQHSRLSSLSGSHTESLPLATSNQVLIKFSAGKLAAPRG 1479
DB 1667 FAFETHALNDVVEHDSQHSRLSSLSGSHTESLPLATSNQVLIKFSAGKLAAPRG 1726
QY 1480 HFVYQAVPRISATQCSSVPEPRYGRKLSGFSVGAIVREFCNSGYALQSPETELPVRG 1539
DB 1727 HFVYQAVPRISATQCSSVPEPRYGRKLSGFSVGAIVREFCNSGYALQSPETELPVRG 1786
QY 1540 ALAOMNVASPTCVPCGGLTERGTILSGPEPYLNSLNCWIKIVPFGAGIOLQVNS 1599
DB 1787 ALAOMNVASPTCVPCGGLTERGTILSGPEPYLNSLNCWIKIVPFGAGIOLQVNS 1846
QY 1600 FVTEQWMSLEVEDGADNTVLMGSPGTTVPALLNSTSNQVLIHFDSDISVSAAGFHE 1659
DB 1847 FVTEQWMSLEVEDGADNTVLMGSPGTTVPALLNSTSNQVLIHFDSDISVSAAGFHE 1906
QY 1660 YKTVGLSSCEEPVAPNSGVATGERLYLVNDVVSQCEBGVALQGHANISCPRGVRRANYP 1719
DB 1907 YKTVGLSSCEEPVAPNSGVATGERLYLVNDVVSQCEBGVALQGHANISCPRGVRRANYP 1966
QY 1720 PPLICACGCTVEMEGVILSPGPGVPPNMCQWIALPVFGAHIQPLNFSTEBND 1779
DB 1967 PPLICACGCTVEMEGVILSPGPGVPPNMCQWIALPVFGAHIQPLNFSTEBND 2026
QY 1780 YBIRNGPYETSRMGRFSSGSELPSSILSTSHETTVYFHSDBQNRGFKLEYQAVELQ 1839
DB 2027 YBIRNGPYETSRMGRFSSGSELPSSILSTSHETTVYFHSDBQNRGFKLEYQAVELQ 2086

```

QY 1840 CPDPEPANGIVRAGAGVNVGQSVTFECTPEGYOITGHPVLTCQHGTRNMDHPLPKCEVPC 1899
Db 2087 CPDPEPANGIVRAGAGVNVGQSVTFECTPEGYOITGHPVLTCQHGTRNMDHPLPKCEVPC 2146
QY 1900 GGNITSSNGTVSPGSPSSODCWMLITVPIGHVRLNLSLOTESGPFITIMGP 1959
Db 2147 GGNITSSNGTVSPGSPSSODCWMLITVPIGHVRLNLSLOTESGPFITIMGP 2206
QY 1960 QGNAPRIVGFTBMAKTVSSSSNOVLKFRHDAATGIFALAFSAVPLTKCPPTILPN 2019
Db 2207 DQNSPOLGVFSGVTAETAYSSITNOVLKFRHDSFNSGCFVILFPAFOLKRCPPRPAPDQ 2266
QY 2020 AAVTNEEFNIDIVRCLPFGTLVGNLITCKLSTLOEGSPRPICEVCPNELLT 2079
Db 2267 ADLITDEDEIDEPKYOCHPEYTLTGSDTLTKLSOLLFQSSPTECAQCPANEVRT 2326
QY 2080 DSGVILSSQSPGSPYQFOTCSMLVNEBDYINISLVEYELSEKQDEFEIFDPSGQSP 2139
Db 2327 ESSGVILSPGYPGNYSQTCAMSIVKKNFNITLFPVDTPOEKQFDALFVFDSSGRSP 2386
QY 2140 LKALSGNYSAPLITVSSNSVYLRSSSDHATNRGFKIRYSAFYCSLPRAPLHGFTIQ 2199
Db 2387 LVLVLSGNTTBOGNSPFRSNHLYLRKSTHATSKGFKIRYAAPYCSLSTLRNGGLNK 2446
QY 2200 TSTQEGSIHFGCNAGYRLVGHSMALCTRHPOGYALMSEALPLCOALSCLPEAPKGMV 2259
Db 2447 TAGAVGSKVHYFCKGPEYRMIGHNATCRNPNFYGVYQWMSAPLCOAVSCGIPAPNGSF 2506
QY 2260 PEKEYTVGTRKAYSGSEGHLOAGAEATACLDITGLMNRNVPPQCVPTCPDVSSISVE 2319
Db 2507 TGNFTLSBKTVTECHGFKLDASQEAATVCOEGLMNRKGPPTCKPVPSPISIEQLSE 2566
QY 2320 HGRMPLIETQYQFOAQLMLICDPYUYTGQVRIRCOANGKSGDSTPCRIISGELP 2379
Db 2567 HYLMLVGSINNEYGAQVILSCSPFYQGRILLQCANNGWSTEEBRPKKALSCSLS 2626
QY 2380 IPRNGHRTGLSYVATATIFSCNGSYTLVGSRVRECNANGIWSGEVRECLAGHGTPEPI 2439
Db 2627 FEPNNGKTLITLITGATVIFTCTNGYTLVGSVRECLANGIWSGEVRECLAGHGTPEPI 2686
QY 2440 VNGHNGENSVYSGSVVYQCNAGFRLIGMSVRIQODHMSGKTPFCVPTTCGHPGNPV 2499
Db 2687 VNGHISGDSGYRDTVVYQCNPRILVGTSTVRICTTSGRGRITVCPVPTTCGHPGNPV 2746
QY 2500 GLTQGNQNLNDVYKVCNPGYMAAGARSQCLASGQWSDMLPTCRIINCTDPEHOENSV 2559
Db 2747 GLTNGTEFNLDVNFCTHGTGRLOGASRAQCRSNQWSSPLPICRVVNSDPSVENAV 2806
QY 2560 ROVASGPHRFSFGTIVSYRCNHGFYLTGPVLSCCODGTWDRPQCLVSCGHPGSP 2619
Db 2807 RHGQGNPESFEYGTSTVWHCKTGFYLLGSSALTCMASGLMDSLRKCLAISGCHGVDA 2866
QY 2620 HSQMSGSYVGVAVRYSCTGKRTLVGNSTRMGLDGHMTGSLPHSGSTSVGCGPGAP 2679
Db 2867 NAVLTGLFTYGAITVQYSCKGQILTGNSSTRVCOBDSHMSGSLPHSGSNGPFCGDPGR 2926
QY 2680 AHGIRLDSDFDPGTVWPFSEAGHVLRGSEERTCOANGSWSGQPCGVISGANGPTBSN 2739
Db 2927 AHGSRILDEPKTSLRFSCEMGHQLRGFAERTCLVANGSMGVOBPCEAVSSGANGPTPN 2986
QY 2740 ARVAFSDGLVSSSIYVECREGYATGLSRHCSVANGTWSGPELTVINCGDPPIANG 2799
Db 2987 GMLISSDGLIFSSSVIYACWEGYKTSGLMTWRHCTANGTWTGAPDCTIISGDPGLTENG 3046
QY 2800 LRLGNDPRVYKTYTQCVGVYMMESHVSVLCTKORTVNGTKPVYCKALMCKPPLIENG 2859
Db 3047 IQGTDTFTFKTVSYQCNPGYLMERPSTPLIRCTKXGTWNGSRPLCKAVLQNPPEVNG 3106
QY 2860 KVVGSDFMNGSSVTVACLEGYQLSLPAVFTCGNSGWTGELPQCPVPCGPGVSSRRR 2919
Db 3107 KVGSGDFRMGASISYSQVQYOLSHALISCGRGVWKGVEYQCLVPCGDPGTPABGR 3166

```

```

QY 2920 EDGFSYSSVSFSCHEPLVYNSPRRRCQSDGTWGTGTPDCTIDPLTTCADPGVQFQI 2979
Db 3167 SGKFTTKSEVFTQCKPPEPLVSSSRRTQADGWGIDOPTCIDPAHTACPDGTPHFGI 3226
QY 2980 QNNSQGVQSVTLFRCKQCYLLQGSTTRTCLPNLTWMSGPPDCVHHKROPTPHNAV 3039
Db 3227 QNNSKGYEVSTVFRCRKHYNHIGSTTRTCLANLTWMSGIQTEBCHACROBETPAHADV 3286
QY 3040 GALDLPBMGYTL 3052
Db 3287 RAIDLPAFGTLLV 3299

RESULT 4
US-10-016-248-46
/ Sequence 46, Application US/10016248
/ Publication No. US20040033491A1
/ GENERAL INFORMATION:
/ APPLICANT: Alabrook et al.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-218
/ CURRENT APPLICATION NUMBER: US/10/016,248
/ PRIOR FILING DATE: 2002-09-20
/ PRIOR APPLICATION NUMBER: 60/254,329
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 60/291,037
/ PRIOR FILING DATE: 2001-05-15
/ PRIOR APPLICATION NUMBER: 60/255,648
/ PRIOR FILING DATE: 2000-12-14
/ PRIOR APPLICATION NUMBER: 60/297,173
/ PRIOR FILING DATE: 2001-06-08
/ PRIOR APPLICATION NUMBER: 60/309,258
/ PRIOR FILING DATE: 2001-07-31
/ PRIOR APPLICATION NUMBER: 60/326,393
/ PRIOR FILING DATE: 2001-10-01
/ PRIOR APPLICATION NUMBER: 60/315,639
/ NUMBER OF SEQ ID NOS: 167
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO: 46
/ LENGTH: 3508
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-016-248-46

Query Match 64.6%; Score 10975; DB 15; Length 3508;
Beet local similarity 62.2%; Pred. No. 0;
Matches 1912; Conservative 453; Mismatches 592; Indels 118; Gaps 11;

QY 1 MAGAPPALL-----LPCLISD-----CCASNQHSVGVGSELYKKQTELKRGV 47
Db 264 ISGTAEASIVLTGMNLSPVYISSKNMLRLHTISDSNHRK-GRNAOFQYKALTELSRGV 322
QY 48 KLPSSKNSQKTSVLTQVGVSGQHNNCPDPGIPERGRKLGSDF-RLGSSVQFTCNEGYDL 106
Db 323 KMLPFSKDSHKNSVLSQGVVALVSHMCLDPGIPENRRARRGSDSRKANVAVQBSCENYVL 382
QY 107 QGSKRTICMVSMPFAMSDHPRVPCARACDAHLRPSGGLTISPNPIYDNNACVWII 166
Db 383 QGSKRTICQAVTTLTAMSDHRTICARTCGSLRGRSPGVITSPVYQYEDNACVWII 442
QY 167 TALNPSKVILAEFEEDLRRGYDTLVGDGQDDQDKTYLYNSQNAQSPSPHTPSRIP 226
Db 443 TTYDDPKVILAEFEELRRGYDTLVGDAGKVGDTIRSLVY-----LTGSSVFD 492
QY 227 ---SMGDDIRQKWTYLAELCRDISSSDARSGSVKRSPTSNABELVAPG---TEIEG 278
Db 493 LIVSMNQWMLH-----LQSDS-----IGSGQFAVYQIEIKG 526
QY 279 SCGDPIPAYGRREGSRFHHGDTLKPECOPAFELVQKALITCKQNNQSAKPGCVFSCF 338
Db 527 GCGDPIPAYGRRTGSSFLHGDITLFECPAFELVGERVITTCQNNQWSGNKPSCVFSCF 586

```

Qy	339	INFTSPBGVLPVNPVDEYGNHLHCWMLILARPSSRTHLANDDVBPQDFELYXOGAT	398
Db	587	FNFTASSGIIILSPNPEEYGNMNCWMLIISEPSRIHLIPNDVDEQDFPLAVKXGI	646
Qy	399	AEAVLFGFSSNOQLPSSITSSGHAARLEFQDHSHTGKGFNITFTFPHNHCPPGPVNV	458
Db	647	SDIIVLGFSSNEVPSQLASSGHIYRLEFQDHSHTGKGFNITFTTGGQNECHPGLPIN	706
Qy	459	GKRFGDSLIQSSSISPLCDEGFLGTQSGEITITCVLKEGSSVNVSAVLRCEAPCCGHLTSP	518
Db	707	GRGRCDFRLLSSSVSPHCDDGFVVTQGSSEITCIIODGNVWSSSTVPRCEAPCGHLTAS	766
Qy	519	SGTILSPWPCGFYDALSACAVIEAOPRYPIKIFTRDKTEVNVDTLEVRGRKTYAPLI	578
Db	767	SGVILPPEWPBYKXDSLHCEMIIEAKPHGISIKMFPDRQTEVNVDTLEVRGPASSBPLI	826
Qy	579	GVYHGTQVPOELISTSNLYLVLFSTDKSHSIGOLFRETITLQSDHCLDPCIIVNOCRH	638
Db	827	GEYHGTQAPQOLISTGNMYLLFTTIDNRSRISGELIHYESHTLESDDLDCIGIIVNGHRH	886
Qy	639	GNDEYVGAIVTFSCDSGYTLLSDGEBLECEPNFOWSRALPSCCALCGEFIGSSGTTILSPG	698
Db	887	GGDFGIRSTVTFSCDPGYTLLSDDEPLVCERNHQNNAHLPCDCLCGGYIQCKSGTVLSPG	946
Qy	699	PPDYPPNNLNTCWTIIETSHGKGVFFTHTFHLESGHYLLITENGSPFOPLRQLTGSRLP	758
Db	947	PPDYPPNSLNTCWTIEVSHGKGVOMIPIFTFHLESHDYLLITTEGSSPSEPARLITGSLP	1006
Qy	759	APIAGLYGNFTAQVRFTSDFSMSYEGENITFSESYDLERCEPEPVPYASIRKGLQFGVD	818
Db	1007	HTIAGLFGNTTAQVRFTISDPSIYEGFNITFSEYDLERCCDPGVAPASRIIGHFVGVD	1066
Qy	819	TLTFSCPEPVBLEGTARITLCGRRRLWSSBLPRCVAECGNSVTGQTLISPNFPVNVN	878
Db	1067	SLTFSCFLGYRLEGATKRLTCLCGGGRVWASVAPLPRCVAECGASVXGNEGTLISPNFPSYD	1126
Qy	879	NNHCIIYSIQPQPKGIOLKARAEBLSGDVLKTYDGNNSARLLGVFSHSEMMGVTLNS	938
Db	1127	NTHCIIYKIEIEAGGHIHLRTRSPQLPEFGDILKTYDGDSSSRPLGTFTKNEILLGLTINS	1186
Qy	939	TSSSLMTDFTDAENTSKGFLHSPILICEPDGCPMKPFYKXHDGHPAGSSVSPSCD	998
Db	1187	TSNHLMEFNTNGSDTDGQFQLYTSPFLVCEPDGCLPNYVYRIRDRBHPFDITVLVLCN	1246
Qy	999	PGYSILRGSEELCLTSGERRTWDRPLPTCVACAGCGTVGSEVSGQVLSPCGPAPRYENLNCI	1058
Db	1247	PGVAMHGSNLTLCISGDRVWDXKPLPCLAIACCGQIHAATSGRLISPOYRAPRYNNLACT	1306
Qy	1059	WTIEAENAGCTIGLHPLVFDEEVDHLRIWDPYBSGVLLKELSGPALPKDLHSTENSIV	1118
Db	1307	WIIEADBGKTIISLHFIVEDTIEAMADILKYWGPVPDSILLKEMSGSALPEDIHSTENSLT	1366
Qy	1119	LOFSTDFPFSOGGALIOFSVSTANSCNDPGIPONGSRSGDSEWEGADSTVFPCCDGYALOG	1178
Db	1367	LOFSDPFISSGSGSIOFSTSIATCNDPQMPQNGTRIGDSREKGDVTFPCDDGYALOG	1426
Qy	1179	SAEISCVKIENRFWMQSPPTPTCIAAPCGDGLTFPSGVILISPMYPPRYPGKCDMKWTVP	1238
Db	1427	QAKITCVOYNNRFWMQDPPTPTCIAACGNLGPAGVILISPMYPPRYPGKCDMKRVKVP	1486
Qy	1239	DYVIALVENIENLEBGYDFLHIYDGRDSLPLIGSFYGSQULPGRIEISSNSLFLAFRSDA	1298
Db	1487	DFVIALIFKSPNMEPSYDFLHIYEGEDSNSPLISYOGSCAPERIESGNSLFLAFRSDA	1546
Qy	1299	SVSNAAGFIDYTEMRESGCPDPSIKXNTRVGSGLKXGSAVTVYCHGYEVBEGSTLSCI	1358
Db	1547	SVGJSGEFLIEFKERKREACCPDGNIMNSTRVGTDPKLGSTITTYCCDSGYKLIDSSITCV	1606
Qy	1359	LGPDGKPPVNNPRPEVCTAPCCGQVYVGSAGVILISPMYPOANTYSGQICLIVFVPPKDYVVF	1418
Db	1607	IGADGKPEMDQVLSNCNAPCCGQYTGSGVILISPMYPHNTYAGQICLXISIVPEKPYVFG	1666
Qy	1419	QPAEFTHALNDVEVHDSHOSRLLSLSGSHTEBSLPLATSNQVLIKESAKGLADPARG	1478

Db	1667	QFAFIFQALNDLAELEFQHTAQAARLLSSGSHGETLLPLATISQIILLRPSAKGSAARG	1726
Qy	1479	FHFYFAQVAPRISATQCSVSEBPRYGRKLSGDSFVGAVIRECNSGVALQGSPEIECLPVP	1538
Db	1227	FHFYQAVAPRISDTQCSVSEBPRYGRIRIGSEBFSAGSIYREECNPGYLLQGSYALHQSVF	1786
Qy	1539	GALAQONVASAPTCVYPCCGNLTERRGILLSPGPEEPLYNLNCYKVIYVEGAGIQIQV	1598
Db	1787	NALAQONMDITPSCVCPSCGNFTQRRGILLSPGPEEPPGNMLNCIMKILIVEGSGIDQPSD	1846
Qy	1599	SFVTEQUMD-SLEVFQDADNTVMILGSPSGTVPALINSTNOUYLFPSYDISVSAAGFH	1657
Db	1847	QFCHGAEILGPEFELHHDGQDVTAPRLSFSGTTVPALINSTNOUYLFQSDISVAAAGFH	1906
Qy	1658	LEYKTVGLSSCEPBAVPSNGVKTERGLVANDVVSFOCEPGYALQGHAIISCMPETVARMN	1717
Db	1907	LEYKTVGLAACQEBALPSNKSIGDRIMANDVLSFOCEPGYTLQGRHISCMPETVARMN	1966
Qy	1718	YPPPLCTACCGGTVEMEVEGYILSPGPGNYPNSNDCSWKIALPVGFAHIOFLNFSTEEN	1777
Db	1967	YPSPLCTACCGGTLSTGGVILSPGPGFSYPPNMLDCTWRISLPLPGYAHIOFLNFSTEAN	2026
Qy	1778	HDYIEINGPSPERSMMGRFSGSLPSLSLSTSHETTIVYHSDSONRPGFKLEUYOAYEL	1837
Db	2027	HDPEIENGPHRTPSMIGQFSQTDLPALLSTHETILHFYSDSONRQCKLAVOAYEL	2086
Qy	1838	QECDEPPEFANGIVRAGAGVYVQGSVTEPCLPGYOLTGHPVLTCQHGTRNRMWDHPLPJCCEV	1897
Db	2087	QNCDEPPEFPQNGMINNDYISVGQSVSEPCFPGYILIGHVPLVTCQHGTRNRMWYFPRCDA	2146
Qy	1898	PCGGNITSSNGTVSPGFPSPYSSSDCVMILTVIGHGYRLNISLTQTEBSGDFITIMD	1957
Db	2147	PCGNVMTSONGTIYSPGFPDEYPLIKDCIMLITVPPGHGYINFTLLQTEVANDYIAVMD	2206
Qy	1958	GPOOTARLGVFRSMKKTVOSSNOVILKFNHDATGCFALAFASAYLTKCPRPITL	2017
Db	2207	GPDDNSPOLGVFSGNTYLETVAYSTNOVLKPHSDFNSNGGFVILNFAPOLKCKQPPRAV	2266
Qy	2018	PNAEVTENEBEFNIGDIVRYRCLBGFPLVGNELITLCKLGYILOECPPIICEVHCPINEL	2077
Db	2267	PQAMLEBDDFEIGDFVKIQCHGYTLVGTDLITLCLSSOLOPBGSLPTCEAOCPANEV	2326
Qy	2078	LTDSTGYILSQSYSPQFQSCMLVREBVDNISLTVEFPSKEQYDEFEIFDGSQ	2137
Db	2327	RTGSSGYILSPGYGYNFNSQTCWSIKVEPNVNIITFVDTFOSEKQFDALLEVFDGSSQ	2386
Qy	2138	SPLLKALISGNYSAPLITSSNSYVLEWSSDHANRKGFKIRYSAPYCSLPRAIDLHPTL	2197
Db	2387	SPLLVVLISGNHTEQNSFTSRSNOLUYLWMSDHAISKGFIRYAAPYCSLTHPLKNGCIL	2446
Qy	2198	GQSTQPGGSIHFPCNAGVRYLVGHSMAICTRHPGYNILMBEAPYLCOALSQGLPEAPRNG	2257
Db	2447	NRTIAGVASKYHIFCKRGYRHWGHSNATTCRRNPLGMYQMOBSLPLCOAVSGGIFESPGNG	2506
Qy	2258	MVFEKEXYTVGKAVYSCSEGYHLQAGAEATAECLDTGMSNRNVPPOCVPTCEBVASSIS	2317
Db	2507	SFTGNEPFLDJSKVUYECHGEFGKLESSQOATVACQEDGLMSNKGKPRMCKCPVACSIAQL	2566
Qy	2318	VEHGRMKLIFETQYQFOAOLMLICDPGYITTTGQAVIRCOANGKWSLGDSTPTCHILISGE	2377
Db	2567	SEHVIWMLVSGSLNEXYGAOVLISCSPPUYLEGMRLLECOANGTWINIDERSPSC-----	2620
Qy	2378	LPFIPNHRIGTILSVGATAIIFSGNSGYTLVGSVHRCMANGLMSGEBVCLAGHGTPE	2437
Db	2621	-----NCHCSPD	2628
Qy	2438	PIYVNGHINGENYSYRGSVVTYQCNAGFRLIGMSVRIICOODHMSGKTPFCVPIITGHBNP	2497
Db	2629	PIYVNGHISGDQFSYRDTIVVYQCNBGFRLVGNYSVRIICQDHMSGQTEVUCVPIITGHBGNP	2688
Qy	2498	VNGLTQGNORPLNDVYKFCVNCPIYMAEGAAARSQCLASGQMSDMLPTCRRIINCTDPHGOEN	2557

Db 2689 AHGFTNGSEFNLNDVNFNTGTYLLQGVSRACGRSNGQSSPLPTCRVNVCSDEGEVEN 2748
Qy 2558 SVRGVASHGPHRSFCTYYSYRCNMGFYLLGTPLVSCQDGTWDRBRPOCLLVSCGHPGS 2617
Db 2749 AIHGGQNFESFEYEMGSLIYHCKGKPHLLGSSALTCMANGLMDSLPKCLAIISGHPGV 2808
Qy 2618 PPHSGSGSYTVGAVVRYSCTGKRTLYVNSSTRMGLDGHMTGSLPHCSGTSVYVCGDPG 2677
Db 2809 PAAVAVLTGELFTYGAIVVHSCRGSESLIGNDRVCOEDSHMSGALPHCTGNNPGFCGDPG 2868
Qy 2678 IPHAGRLDSDPEGTVMRPSCEAHVLRGSSERTCOANGSVSGSQPECGVISCNPSTP 2737
Db 2869 TPAHGSRLDDDFKTSKSLRFSCEMGHQLRGSEPERCTCLANGSNGGLPVCBAVSCGNPCTP 2928
Qy 2738 SNARVVSFSDGVFSSSIYVECEGEGYATGLSRHCSVNGTWSGPECLVINCDDPGIPA 2797
Db 2929 TNGMIVSSDGLIFSSSVIYACMEGYKTSGLMTRHCLANGTWGTAPDCTIISGDPGTLA 2988
Qy 2798 NGLRLGNDPRYKVTYVQCPGVMMESHRSVLSCTKDRTNWNGTYPVCALMKCPPLIP 2857
Db 2989 NGIOGTDFTFNKTVSYQCNPGYVMEAVTSATIRCTKGRNNPSKPVCAVLCTQPPVQ 3048
Qy 2858 NGKVVSGDPMGSSVTVYACLEGYQLSLPAVFTCEGNSWTBELPQCFPVFCGDPVSRG 2917
Db 3049 NCTVSGSDPRWSSSISYSCMDGYQLSHSAILSCERGVWKGEIPQCLPVFCGDPGIPAG 3108
Qy 2918 RREDGFSYRSVSFSCHPPLVYSGPRFCOSDPTWSGQPSCIDPLTTLTCDPGRVPOF 2977
Db 3109 RLSGSFYTKSEVFPQCKSPFLIVSSSRVCOADGTMSGIOTPIDPAHNTCPDPGPHF 3168
Qy 2978 GIQNSQGYVQVSTYLFRCQKGYLLQGSTTTCPLNLTWSGTPDVCVPHHCROPEPTHA 3037
Db 3169 GIQNSRQYEVGSLTYFFRCRKGYHIQGSTTRCTLANLTWSGIQECIPHACROPEPTAHA 3228
Qy 3038 NVGALDLPMSGYTLI 3052
Db 3229 DVRAIDLPFTGYTLV 3243

RESULT 5
US-10-016-248-47
Sequence 47, Application US/10016248
Publication No. US20040033491A1
GENERAL INFORMATION:
APPLICANT: Alabrook et al.
TITLE OR INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-218
CURRENT APPLICATION NUMBER: US/10/016,248
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: 60/254,329
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/291,037
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/255,648
PRIOR FILING DATE: 2000-12-14
PRIOR APPLICATION NUMBER: 60/297,173
PRIOR FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: 60/309,258
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/326,393
PRIOR FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: 60/315,639
PRIOR FILING DATE: 2001-08-29
NUMBER OF SEQ ID NOS: 167
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 47
LENGTH: 3389
TYPE: PRT
ORGANISM: Homo sapiens
US-10-016-248-47

Query Match Best Local Similarity 61.7%; Score 10475.5; DB 15; Length 3389;
59.9%; Pred. No. 0;

Matches 1843; Conservative 430; Mismatches 565; Indels 237; Gaps 12;

Qy 1 MAGAPPALL-----LPCLISD-----CASQHRSHVGVGSELYKQIELKRGV 47
Db 264 ISGTEASVILWTGMNLPSPVITSSKNWLRHFTSDSNHRRK-GFNAPQVYKAIELKSRGV 322
Qy 48 KLPMSKDNQSQTSLTVQGVSGHNNCPDGGIPEKRGKSGSP-RIGSSVOTFCNMGYL 106
Db 323 KMLPSKDGSHKNSVLSQGGVALVSHHCLDPGIPENGRRAQSDPSRGAVANVPSCEDNYL 382
Qy 107 QGSKRTCKMVSIMFAMSDDRVCARNCDALHKGPSGITSPNFIQYDNNACVWII 166
Db 383 QGSKSITCORVETTLAAMSDHRPICARFCGSLNKGPSGVITSPNVPQYEDNALCVWII 442
Qy 167 TALNPSKVIKLAFEEDFLERGYDTLTVGGGDDGQOKTLYVMSQACSDSPHTPSRIPE 226
Db 443 TTTPDPKVIKLAFEEDFLERGYDTLTVGGAGVGRSVLVV-----LTGSSVVD 492
Qy 227 ---SNSGDIWRQKWTVLEICRDISSDARSQSVRSKPSYTSNAVELVAPG----TEIRQG 278
Db 493 LIVSMNQWMLH-----LQSDDS-----IGSPFRAVYGEIEKG 526
Qy 279 SCGDGPIPAYGRREGSRFHGDTLKFECQAPAFELYGOKALITCQKNQMSAKKPCGVFSCF 338
Db 527 GCGDPEIIPAYGARTSSFLHGDTLTFECPAPELVEGEVITCOQNNQMSGNKPCVFCF 586
Qy 339 FNFPSGVLSPNYPEDYGNHLHCVWILARPESRHLAFNDIDVEQDFPLVYKQAT 398
Db 587 FNFASSGIILSPNYPEYGNMNCVWIIIEPGRHILINDDVEQDFPLVYKQAT 646
Qy 399 AAPVLGFSGNOLPSSITSSGHVAREFQDHSITGKRGFNITTFPHNCPDGPVNV 458
Db 647 SDITVLGFSGNEVVSQASSGHIYRLEFQDHSITGGEFNITTFEQNCCHDGPAPIN 706
Qy 459 GRRFGDSIQLGSSISFLGDEGFLGTGQSEITTCVYKGSVYVMSAVVACEAPCCGHLTSP 518
Db 707 GRRFGDRFLGSSSVSFHDDDFVKTQGSSEITTCVYKGSVYVMSAVVACEAPCCGHLTAS 766
Qy 519 SGTIISPCWPGFYDALSANVIAOPQYPIKITFDRFKTEVNYDTLEVRDGRYSAPI 578
Db 767 SGVILPQWPGYDYSLHCWEIIEAKPGHSIKMTFDRFQREVVNYDTLEVRDGRASSPLI 826
Qy 579 GVYHCTQVQPLISTSVYLLFSTDKSHSDIGQALYETITLSDHCLDPGIPVNCQRH 638
Db 827 GEYHCTQVQPLISTGNFMVLLFTDNRSSIGLHYEVSUTLESDBCGIPVNVGHRH 886
Qy 639 GNDPYVALVTFSCDSGYTLSDGEPLCEBNFQMSRALPSCBALCGGFIQSSGTTISP 698
Db 887 GGDGGRSTVTFSCDPGYTLSDDEPLVCERNHQNHALPSCDALCGYIQKSGTYLSPG 946
Qy 699 FPDYPPNNLCTWIIETSHGKGVFTFTHLESCHDYLLITNGSFTQPLRQLTGRRLP 758
Db 947 FPDYPPNLSLCTWITEVSHGKGVMTFTHLESCHDYLLITDGSFSEPARLTSVLP 1006
Qy 759 APISAGLYGNFTAOVRPISDFSMSEYGFNTFSEYDLEPCEBEVAVAYSIRKLOCGYVD 818
Db 1007 HTIKAGLFGNFTQRLARISDFSIYEGFNITSEYDLEPCCDGVAFSRRIFHGVGD 1066
Qy 819 TLTFSCTPGYRLAGTAITLGLGRRRLMSSPLPRCAVACGNSVYTGQTLSPNFPVNY 878
Db 1067 SLTFSCTGLGRLGATKLTCLGGGRVWMSAPLPRCAVACGASVYKNEGTLSPNFPVND 1126
Qy 879 NNHECTYSIQTPGKGIOLKARAFELSEBDVLYKVYGNNNNSARLGLVESHSEMGTYLNS 938
Db 1127 NTHECTYKIETENGKGIHLARTSFQLEBDTLYKVDGKSSSRPLGTFTKMLLGLILNS 1186
Qy 939 TSSSLWLDFTTDANTSKEFELHFSFELIKCEDGTPEFGYKVDHDEGHFAGSSVSFSCD 998
Db 1187 TSNHMLWLEFVINGSDDTQGFQLYTYSFDLVKCEDGIPRYGRIIRDBGHFTTVVLVXSN 1246
Qy 999 PGYSLRGSBELCLSGERTWRPLPTCYAECGGTVRGEVSGQVLSPGYPADYENHNLCT 1058
Db 1247 PGYAMHGSNTLTCLSGDRRWMDKPLPSCTAECGQIHAATSGRIILSPGYAPADYDNNLHCT 1306

QY 1059 WTIEBAGCTTIGLHFLVPTDEEVNDVLRIMDGPVSGVLLKELSGPALPKDILSTFNSVY 1118
 DB 1307 WIEADBGKTIISLHFLVPTDEEMADHILKWDGPDSDILKEMSGSLPEDIHSTFNSLT 1366
 QY 1119 LQSTDEFTSKOGFAIOFVSSTATSCNDPGIPONGSRGSDSWEAGDSTVEQCDPGVALQ 1178
 DB 1367 LQSDDFEISKSGSIOFSTSIATCNDPGMPONGTGRSREAGDVTQCDPGVLOQ 1426
 QY 1179 SAEISCKIENRPFWQSPPTCIAPCCGDLTGPEGVILSPNYPEPIYPGKECMKVTVP 1238
 DB 1427 QAKITCOLNNRPFQWQDPPTCIACCGGNLTGPAGVILSPNYPPQYPGKECDRVKVP 1486
 QY 1239 DVIYALFENIENLEPGDPLHIYDGRDLSPLGSEFGSOLPGRIEISSSLFAPRSDA 1298
 DB 1487 DFIYALIFKSPFNEBSYDFLHIYEGEDSNSPLISYQSOAPEKIESGSLFAPRSDA 1546
 QY 1299 SVSNAGFVIDYTENPRESCFDPGSIKNGTRVSGDLKSGSVTYXCHGVEGTSLSCI 1358
 DB 1547 SVGLSGFAIEFKERREACFDPGIMNGTIVGDFKLGSTITTYOCDGKYKILDESSITCV 1606
 QY 1359 LGPDGKPVNNPRPVCTAPCCGGQYVSGDVVLSFNPYQNTYSGQICLYPTVPKDYVFG 1418
 DB 1607 IGADGKPSMDQVLPSCNAPCGGQYTGSEGVLSFNPYHNTAQOICLYSITVPKDFVFG 1666
 QY 1419 QPAFPHALNDVYEVHGHGSHSLSLSGSHTEGSELPLATSNQVLIKESAKLAPARG 1478
 DB 1667 QPAFPHALNDVLEFPGTHAQARLLSLSGSHGELPLATSNQVLIKESAKLAPARG 1726
 QY 1479 FHEVYQAVPRTSATQCSSVEPEPRYKRLGSDFSVGAIVRECSNGVALQSPETECUPV 1538
 DB 1727 FHEVYQAVPRTSATQCSSVEPEPRYKRLGSDFSVGAIVRECSNGVALQSPETECUPV 1786
 QY 1539 GALAQNNVAPTCVPCGGNLTERRGITLSPGPEPYLNSLNCWKIVDEGAGIQIYV 1598
 DB 1787 NALAQNNVAPTCVPCGGNLTERRGITLSPGPEPYLNSLNCWKIVDEGAGIQIYV 1846
 QY 1599 SFYTEQWMD-SLEVPDQADNTVMLGSPGCTTPALNSTSNQIYHAFYSDISVSAFGH 1657
 DB 1847 QFCHGABLGLPFETHDGDVTAAPRLGSPGCTTPALNSTSNQIYHAFYSDISVSAFGH 1906
 QY 1658 LEYKTVGLSSCPREPAVNSNGVKTGERLYLVNDVVSFOCEPGVALOGHAHISCMPTVARM 1717
 DB 1907 LEYKTVGLSSCPREPAVNSNGVKTGERLYLVNDVVSFOCEPGVALOGHAHISCMPTVARM 1966
 QY 1718 YPPEPLCIAQCGGVEENEGVILSPGPGNYPNSNDCSMKIALPVFGAHIQFINFSTEPN 1777
 DB 1967 YPPEPLCIAQCGGVEENEGVILSPGPGNYPNSNDCSMKIALPVFGAHIQFINFSTEPN 2026
 QY 1778 HDYIEINGBYETSRMMGRFSGSELBSLSLSTHETTVYFRSHDSQNRPGFKLEYAYEL 1837
 DB 2027 HDYIEINGBYETSRMMGRFSGSELBSLSLSTHETTVYFRSHDSQNRPGFKLEYAYEL 2086
 QY 1838 QECDDPEPPANGIVRAGAGVYVQGVTECPGLGYOLTGHPVLTQCHGNRMMDHLPKCEV 1897
 DB 2087 QECDDPEPPANGIVRAGAGVYVQGVTECPGLGYOLTGHPVLTQCHGNRMMDHLPKCEV 2146
 QY 1898 PCGNNITSSNGTVSPGFPSPBYSSODCWLITVPIGHGVRNLISLLQTEPSGDFITIM 1957
 DB 2147 PCGNNITSSNGTVSPGFPSPBYSSODCWLITVPIGHGVRNLISLLQTEPSGDFITIM 2206
 QY 1958 GPOQTARLGVTFRSMKATVQSSNOVLKFRHDAATGGIFALAFABAPLTCKPPTIL 2017
 DB 2207 GPOQTARLGVTFRSMKATVQSSNOVLKFRHDAATGGIFALAFABAPLTCKPPTIL 2266
 QY 2018 PNAEVVENEENEFNGIIVRYACLPGLTIVGNEILITCKLGTVLOEGEPPICEVCHPTNEL 2077
 DB 2267 PNAEVVENEENEFNGIIVRYACLPGLTIVGNEILITCKLGTVLOEGEPPICEVCHPTNEL 2326
 QY 2078 LDTSTGVILSQSYGYSYPOFQTCMLVREBPNYISLTVEYFSEKQYDEFEIFDGSQ 2137
 DB 2327 RTGSSGVILSPGYGNFNSQTCMSIKVEBNYITITFVDTPOSEKQFDALLEVDDGSSQ 2386

QY 2138 SPLKALSGNYSAPLIVTSSNSNVYLRWSSDHAYNRKFXIRYRAPSCLPAPRLHGFIL 2197
 DB 2387 SPLVVLDSGNHTQSNFTSNQDYLKRWSTDHATSKKGFIRAPRCSLTHPLKNGIL 2446
 QY 2198 GQSTQPGGSIHFCNAGYRLVGHSMALCTRHPOGYHLMSEALPLCOALSGLPEAPKNG 2257
 DB 2447 NRTAGAVGSKVHYFCKRGYVMGHNSNATCRNPLGMYQMDSLPLCOAVCSGIPESPGNG 2506
 QY 2258 MWGKEYTVCATVYKAVYCSSEGYHQAQAEALAECLDGMENRNVPPQCVVTPBDVSSIS 2317
 DB 2507 SFTGNEFTLDSKVYVECHBEFKLESSQQAATAVQOEGLMSNKKKPPWCKRVACPSIEAOL 2566
 QY 2318 VEHGRWRLPETYOFOAOLMLICDPXYXTYTGQVIRCOANGKMSLSDSTPTCRIISCGE 2377
 DB 2567 SEHVITRLVGSILNEGAQVLLSCSPGYIEGRRLRCOANGKMNIGDERPSR----- 2620
 QY 2378 LPIPNGRHIGLTVYVQATAFSCNSGYTLVGSRVBECNANGLSMGSEVRCLAGHCGTPE 2437
 DB 2621 -----AGHCGSPD 2628
 QY 2438 PIVNGHNGENYRSGSVYVQCNAGRLLJMSVRICQODHNSGKTPFCVPIITCGHPRNP 2497
 DB 2629 PIVNGHISGDFSGYRDTVVYVQCNPGFRLVGTSVRICLODHKMSGOTPVCPVITCGHPRNP 2688
 QY 2498 VNGLTQGNOLNDVYKFCVCPNPGYMAEGARSOCCLASGQSDMLPTCRITINCTDPHQEN 2557
 DB 2689 AHGFTNGSEPNLNDVNFCTNTGYLLQVSRACQCRSNGQSSPLPTCRVNCSDPGIVEN 2748
 QY 2558 SVROVHASGPHRPSFGTTVSYRCNNGFYLLGTPEVLSQCGDGTWDRPPOCLVSGHPS 2617
 DB 2749 AIRHQGNPFESFEYEGMSILYHCKKGPHLLGSALUCMANGLMDRSLPKCLALSCHPGV 2808
 QY 2618 PPHSQNSGDSYTYGAVYRSCIGKRLTVNGSTTMCGLDGWTSLPHRCSTSVGCGDPG 2677
 DB 2809 PAAVAVLTEGELFTYGAVVHYSCRSSESLIGNDTVECOEDSHMSGALPHCTNNPGFCGDPG 2868
 QY 2678 IPAAGIRLGDSPDGTVMRPSCEAGHYLRSSERTCOANGSMGSGOPEGVISCQNGTPT 2737
 DB 2869 TPAHGRHLDGDFTKSLRSCMGHQLKRSPPRTCLINGSMGLO----- 2914
 QY 2738 SNARVYFSDGLVFSSSIVYECREGYVATGLLSRHCNSVNGTWGSDDECLVINGCDGIPA 2797
 DB 2915 ----- 2914
 QY 2798 NGRLRNDPFRKNTVYQCVPGTMESHVSVLSTKDRITWNTKPYCKALMKKPPILP 2857
 DB 2915 -----PVCEAVLCPPOPPVQ 2929
 QY 2858 NGKRVGSDPEWGSVYACLEGYOLSLPAVFTCEGNGSWTGELOPCGPPVCGDPGVSRG 2917
 DB 2930 NGTVEGSDPEWGSVYACLEGYOLSLPAVFTCEGNGSWTGELOPCGPPVCGDPGVSRG 2969
 QY 2918 RREDRGFSYRSVYFSCHPRLVYVGSBPRFCQSDGTWGTQSPICIDPTLTTCADPGVPOF 2977
 DB 2990 RLKGSFTYKSEVYFPOCKSPFLLVGSRRVCOADGTWGTQSPICIDPTLTTCADPGVPOF 3049
 QY 2978 GIONNSQIOVGSVTLFRQCKGYLLQSTTRTCLPMLTWSGTTPDCVPHHCROPEPTTHA 3037
 DB 3050 GIONNSRKYEVGSVTFVFRCKGYHIOSTTRTCLAMLTWGTGOTECIPHACROPEPTTHA 3109
 QY 3038 NVGALDLPMSGYTLI 3052
 DB 3110 DVRAIDLPRTGYTLV 3124

RESULT 6
 US-10-408-765A-2286
 ; Sequence 2286, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Boia D.
 ; APPLICANT: Zhang, Bing

/ APPLICANT: Gibson, Bradford W.
 / APPLICANT: Taylor, Steven W.
 / APPLICANT: Glenn, Gary M.
 / APPLICANT: Warnock, Dale E.
 / TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 / TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 / FILE REFERENCE: 660088.465
 / CURRENT APPLICATION NUMBER: US/10/408,765A
 / NUMBER OF SEQ ID NOS: 3077
 / SOFTWARE: FASTSEQ for Windows Version 4.0
 / SEQ ID NO 2286
 / LENGTH: 3389
 / TYPE: PRF
 / ORGANISM: Homo sapiens
 / US-10-408-765A-2286

Query Match 61.7%; Score 10475.5; DB 16; Length 3389;
 Best Local Similarity 59.9%; Pred. No. 0;
 Matches 1843; Conservative 430; Mismatches 565; Indels 237; Gaps 12;

1 MAGAPPALL-----LPCSLISD-----CCASNOHSHVGVSELYKQILKSRGV 47
 264 IGTAPRSIWLGMPLSPVLISSKWLRLHFTSDSNHRRK-GFNAQFVKKXILKSRGV 322
 48 KLPSPKDSQKTSVLTQVGVSGHNMCPDPIPERKRLGSDF-RLGSSVQFTNEGVDL 106
 322 KMLPKSDSHKNSVLSQGVVALVSHMCLDPIPERGRAGSDPFRVGNANVOFSCEDNYL 382
 107 GSKRITOMKYSDMAASDHRPVCRAMCAHLRSPGITSNPFLOYDNNAHCWII 166
 303 QOSKITTQRTVETLAAWSDHRPICRATCGSNLRGSGVITSPNYEVQEDNAHCWVI 442
 167 TALNPKYIKLAFEEFDLERYDITLVGDGGDDGQDKTVLYMSQNASDDBPHPGSIPE 226
 443 TTTDPDKYIKLAFEEFDLERYDITLVGDAGKVDITSLVY-----LTGSSVPD 492
 227 ---SMSGIWROKMTVLEICRDISSDARSVSKSPKTSNAVELVAPG-----TEIEQG 278
 493 LIVSNQNMWLNH-----LQSDDS-----IGSPGKAVYQIEBK 526
 279 SCGDPGIPAYGRGSRFHGDTLKFEQOPAFELVQKAITCQKNQMSAKRGCVPSCF 338
 527 GCGPBGIPAYGKRGTSSFLHGDITLFECPANFELVGRVITCCQNNQMSGNKSCVPSCF 586
 339 FNFTSPGCVLSPVPEYDGNLHCWMLIARPESTRHLAFNDIDVEPODFLVIQDGT 398
 587 FNFTASGIIILSPVPEYDGNMNCWMLISEPDSRIHLFNDPVBPODFLAVDNGI 646
 399 AEAIVLGFSGNOLPSSITSSGHVAREFOTDSTGKRGNTITFTFRHNECPDGPVPA 458
 647 SDITVLITFGSNEVPQSALSSGHVLEFQSDHSTTGRGNITVYTFQNECHDPIPIFN 706
 459 GKRFDSLOQSSISFLCDEGFLATQGSFETICVLKESGVNMSAVLARCEAPCGHLS 518
 707 GRFRGDFLLGSSVFFHCDGPFVKTQGSSEITCIIIDGNVMSSTVPRCEAPCGHLS 766
 519 SGITLSPGMEGFYKDALSCAWLEAOPGYPIKITEPRFTEVNYDTLLEVRDRTYSAPLI 578
 767 SGVILPBGWPGYKDLHCEWIIIEARGHSHKMTFDRFQTEVNYDTLLEVRDPAASSPLI 826
 579 GVVHGTQVPOFLISTSNVYLFLSTDKSHSDIGFOLARETITLQSHCHLCPITPNQGRH 638
 827 GEVHGQAPQFLISTGNFMVILFTTNSASSIGFLIHYSVTLSESSCLDPIPIVNGRKH 886
 639 GNPVYCALVTRSCDGYTLSDGEPLCEPNFQMSRALPSCALCGGFIQSSGTTILSPG 698
 887 GGDFGIRSTVTFSCDGYTLSDGEPLVCEBNQNMALPSCALCGGYIQSSGTTILSPG 946
 699 FPDFYNNINCTWIIITSHGKGFFTFHFEHLESGLDYLITLBNSSFTQPLRQLTGSFLP 758
 947 FPDFYNNINCTWIIITSHGKGFFTFHFEHLESGLDYLITLBNSSFTQPLRQLTGSFLP 1006

759 APISAGLYGNFTAOVRFISDFSMSEYGNITSEVLEPCEPEVPAYSIRKLOQGVCD 818
 1007 HTIKAGLFGNFTAOVRFISDFSMSEYGNITSEVLEPCEPEVPAYSIRKLOQGVCD 1066
 819 TLTPSCPGYRLGRTARITCLGRRRLMSSPLPRCAECGNSVTQGTLLSPNFPVNN 878
 1067 SLTPSCPGYRLGRTARITCLGRRRLMSSPLPRCAECGNSVTQGTLLSPNFPVNN 1126
 879 NNEHCYSIOTOPKIOKAPAFELSEBDVLYKVDGNNSARLLGVSHSEMGTLLNS 938
 1127 NNEHCYSIOTOPKIOKAPAFELSEBDVLYKVDGNNSARLLGVSHSEMGTLLNS 1186
 939 TSSSLMLDFTDAENTSKGFEHFSFELIKCEDPTPKFYKXNDEGHFASSVSFSCD 998
 1187 TSNHMLWFENTNGSDTDQGFQTLTYSFDLVKCEDDGIPIVYGRIRNDEGHFTTVLYSCN 1246
 999 PGYSLRGSEBELCLSGERTWDRPLPTCAECGTVRGESQVILSPGYAPAEHNLCTI 1058
 1247 PGYAMHGSNTLTLCLSGDRVWMDKPLPSCIAECGQIHAATSGRIILSPGYAPAEHNLCTI 1306
 1059 WTIEAGCTIGLHFLVPTTEVDVLRIMQDPVBSGVILKELSGPALPKDLSFENSIV 1118
 1307 WTIEADPGKTIHLHFTVPTTEMAHDLIKYMDQPDVSDILKEMSGSALPEDIHSTFNSLT 1366
 1119 LQSTDFFTSKGPAIORSVSTATSCNDPGLIPQNSRGSDSWEAGDSTVFOCDPQYALOG 1178
 1367 LQFDDDFLTKSGFISQESTSIATCNCPMPQNGTRVGDREAGDVTVFOCDPQYALOG 1426
 1179 SAEISCVKIENFFPQBPPTCIAFCGGDLGPGSVILSPVPEYPRGKCDMKVTVSP 1238
 1427 QAKITCVQANRFFPQBPPTCIAFCGGDLGPGSVILSPVPEYPRGKCDMKVTVSP 1486
 1239 DYVIALVFNIPLEBYDPLHIYDGRDLSPLISFYGSQULPGRIESSNGLFLAFPSDA 1298
 1487 DVIILIFKSNMEBSYDFLHIYGEDENSPLISYQSSQAPBEIESSNGLFLAFPSDA 1546
 1299 SVSNAGFIIDTENDRESCPDPSIKNGTRVGSGLKIGSSVTVYCHGQYEVETSTLSCTI 1358
 1547 SVGLSGFIIRKREKREACFPDPSIKNGTRVGSGLKIGSSVTVYCHGQYEVETSTLSCTI 1606
 1359 LGPDGKPVNPNRPVCTAPCGGYVSGDGVLSBNYPONTYSAGICLPTVPVDPDVPVFG 1418
 1607 IGADGKPSMDQVLPSCNAPCGGYVSGDGVLSBNYPONTYSAGICLPTVPVDPDVPVFG 1666
 1419 QFAFETALNDVAVNHGSHSRLSLSGSHTESLPLATSNQVILIKESAKGLAPARG 1478
 1667 QFAFETALNDVAVNHGSHSRLSLSGSHTESLPLATSNQVILIKESAKGLAPARG 1726
 1479 FHFYQAVPRITATQSSVPEPRYGRKLGSDFSVGAIVREPCNSGYALQSPETELCPVP 1538
 1727 FHFYQAVPRITATQSSVPEPRYGRKLGSDFSVGAIVREPCNSGYALQSPETELCPVP 1786
 1539 GALQNNVSAPTCVPCGSLTERRGTTILSPGPEPYLNSLNCWIKIVPEGAGIOIQUV 1598
 1787 NALQNNVSAPTCVPCGSLTERRGTTILSPGPEPYLNSLNCWIKIVPEGAGIOIQUV 1846
 1599 SFVTEQWMD-SLEVPQADNTVMIGSFQTTYPALNISTSNLYLHFTSDISVSAAGRH 1657
 1847 QFCGAGLGLPFEIHGSDVTAARLPSFGSTVYPALNISTSNLYLHFTSDISVSAAGRH 1906
 1658 LEVYTVLSSCPPEAVPNSGVKTGERVLYVNDVVSFOCEEGVALOGHAHISCMPTGYARRNN 1717
 1907 LEVYTVLSSCPPEAVPNSGVKTGERVLYVNDVVSFOCEEGVALOGHAHISCMPTGYARRNN 1966
 1718 YPPEPLCIAOCGGVEEMEGYILSPGPGYVPSNMDCSWKIALPVGGAHIOFLNFTSEBN 1777
 1967 YPPEPLCIAOCGGVEEMEGYILSPGPGYVPSNMDCSWKIALPVGGAHIOFLNFTSEBN 2026
 1778 HDVIEIRNGEYETSRMMGRFSGSELPSSLSHETTYFHSDBHONRGEFLAEQAYEL 1837
 2027 HDVIEIRNGEYETSRMMGRFSGSELPSSLSHETTYFHSDBHONRGEFLAEQAYEL 2086
 1838 QBCPDPEPRANGIVRGAGYVNGSVTFECLPGYOLTGHPVLTLCQHGTRNMDHPLPKCEV 1897

Db 2087 QNCPDPPFQNGWINDSDYSGSVSEFCYPGYILIHAPVLTOQHGRINRWNPFRPCDA 2146
Qy 1898 PCGGNTSSNGTVSPSPSPSSQCCWMLITVPIGHVRLNLSLQTEPSGDFITWD 1957
Db 2147 PCGGNTSSNGTVSPSPSPSILKDCIMLITVPPGHGVINFTLLQTEAVNDYIAVMD 2206
Qy 1958 GPGOTARLGVFRSMKKTVOSSNOVLKKEHNDATGIFALAFSAAYLTKCPRTIL 2017
Db 2207 GPDONSQOLGVFSGNALLETAVSTNOVLKPHSDFSNGGFVNLNFAFOLKCKQPPAV 2266
Qy 2018 PNAEVVTESEFNIGDIYRYRCLPFTLVGNEILTKLGTLYLOFEPPICEVHCPTNEL 2077
Db 2267 PQAMELDEDDFEIGDFVYKQCHGYTLVGTDLTKLSSQLQREGSLPTCEAQNANV 2226
Qy 2078 LTBSTGYLSQSYGSPQFOTCSWLYRVEDPNVLSLTFEYSEKQYDEBEIFDPSGQ 2137
Db 2127 RTGSSGVILSPGYGPNFNSQTCMSIKVEPNVITTFVDTFOSEKQFDALLEVFGSSGQ 2386
Qy 2138 SPLKALSGNYSAPLITSSNSVYLWMSDPAHNRGFKRYGAPYCSLPRALHGFIL 2197
Db 2387 SPLVLVSGNHTGOSNFTSRNQLYLRWSTDHATSKGFKRYAAPYCSLHPKNGCIL 2446
Qy 2198 GOSTOPGSIHFGCNAGYRLVGHSMALCTRHPOGYHLMSEAPLPCQALSCGLPEAPKNG 2257
Db 2447 NRTAGAVSKVHYCKRPGYRNWGHSMATCRNPLGMQOMDSLTPLCQAVCGIDPESKNG 2506
Qy 2258 MVEKEVTVGTKAVYSCSEGYHQAABATACLDGTGLMSNRNVPPOCVPTCPDVSIS 2317
Db 2507 SFTGNETFLDSKVYECHEGFKLESSQOATVACEDGLMSKGRPMCKPACPSIABOL 2566
Qy 2318 VEHNRMLIFETQYQFOAQLMLICDDPGYVYTGQVIRIQANGKSLDSTPTGIIISGE 2377
Db 2567 SEHVMWLVGSLWEYGAQVLLSCSPGYLEGMWLLRCQANGTWNIIDERSKSC----- 2620
Qy 2378 LPIPPNGHRICTLSVYGATAIFSCNSGYTLVGHVRECMANGLWMSGESEVCLAGHCSTPE 2437
Db 2621 -----AGHCSPD 2628
Qy 2438 PIYNGHNGENYSYRGSVYVYQCNAGFRLLGMSVRIICQODHMSGKTPPCVPITCGHPNP 2497
Db 2629 PIYNGHISGDFSRDPTVYVQCNNGFRLVGTSVIRICLDHMSGQTPVCVPTICGHGPNP 2688
Qy 2498 VNGLTQONPFLNVVKEVNCNPGMAGAAASQCLASGMSDMPTORINICNDPQHON 2557
Db 2689 AHGFTNSEEFLNVNFTCTGYLLQGVSAQCRSNGWSSPLPTRVYVNCSDPGVEN 2748
Qy 2558 SVRQVHASGPHRFSGFTTVSYRCHGFLYLTGTPVLSQCGDGTWDPRPPOCLVSCGHPS 2617
Db 2749 AIRHGQONFESFEYGMSTLYHCKKGRHLGSSALTCMANGMLNRSPLKCLATISGHPGV 2808
Qy 2618 PPHSQMSGDSYTVGAVVRYSCIGRRLVGNSTRMCGLDGHWGSLPHSCGTSVGVCDPG 2677
Db 2809 PANAVLITGELPTYGAVVHYSCRGSESLIGNDTRVCOEDSHWSGALPHCTGNNPFGCDPG 2868
Qy 2678 IPANGILGDSFDDCTWNRSCGAGHYLRGSSERTCOANGMSGSSQCEGYSICGNPCTP 2737
Db 2869 TPAHSGRLGDFKTKSLRFSCGEMHOLRGSPEPTCLLNGWSGLQ----- 2914
Qy 2738 SNARVFSDDLVSFSSIYVECRGEGYATGLLRHCSVNGTGTGSDPCLVINCDDPGIPA 2797
Db 2915 ----- 2914
Qy 2798 NGRLGNDPFRYKNTVYQCPGYMESHRSVLSCTYDRITWNGTKPVCKALMCKRPPLIP 2857
Db 2915 -----PVCEAVLCPQPPVQ 2929
Qy 2958 NGKVVSGDFPMWGSVYTALEGYQLSLPAVFTCEGNSWTGELPOCFPVFGDGPVSRG 2917
Db 2930 NGTEGSDFRWGSISYSQMDGYQLSHALISCBGRGWKEISIPQCLPVFGDGIABEG 2989
Qy 2918 RREDGFSYSSVSFSCHPPLVLVSGPPRRFCQSDGTWGTGTOPSCIDPLTTCADPGVPOF 2977

Db 2990 RLSCGSFTYKSEVFFOCKSEFLLVWSSRRVQCADGTWSGIQPTCIDPAHNTCPDPTPHR 3049
Qy 2978 GIONNOGYGVSTVLPPOCKGYLLQGSTTRTCLPVLTSWGPDPDVPRHCRQETPTH 3037
Db 3050 GIONSSRGYGVSTVFPCKRGYHIQSTTRCLANLWTSIGTQECTPHACRQETPAHA 3109
Qy 3038 NVGALDLPMSGYTLI 3052
Db 3110 DVRAIDLPTGYTLV 3124

RESULT 7
US-10-467-433-18
; Sequence 18, Application US/10467433
; Publication No US20040087773A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION;
; APPLICANT: LAL, Preeti G.; BAUGHN, Mariah R.;
; APPLICANT: YAO, Monique G.; CHAWLA, Narinder K.;
; APPLICANT: ELLIOTT, Vicki S.; XU, Yuming;
; APPLICANT: HONCHELL, Cynthia D.; YUE, Henry;
; APPLICANT: DING, Li; GIETZEN, Kimberly J.;
; APPLICANT: ISON, Craig H.; LU, Dyanth Alina M.;
; APPLICANT: HAFALIA, April J.A.; GANDHI, Ameena R.;
; APPLICANT: THANGAVELU, Kavitha; SANJANMALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; RAMKUMAR, Jayalaxmi;
; APPLICANT: GRIFFIN, Jennifer A.; SMARNAKAR, Anita;
; APPLICANT: AZIMZAI, Yaida; SAPPERSTEIN, Stephanie K.;
; APPLICANT: BURFORD, Neil; LEE, Ernestine A.;
; APPLICANT: LU, Yan; TRAN, Uyen K.;
; APPLICANT: MAROUS, Joseph P.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PR-0899 USN
; CURRENT APPLICATION NUMBER: US/10/467,433
; PRIOR APPLICATION NUMBER: PCT/US02/03709
; PRIOR FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: PCT/US02/03709
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/266,117
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/269,618
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/271,118
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,486
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/274,436
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/334,229
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/353,284
; PRIOR FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 18
; LENGTH: 1958
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 55061615CD1
US-10-467-433-18

Query Match 53.1%; Score 9025.5; DB 15; Length 1958;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 1692; Conservative 3; Mismatches 6; Indels 257; Gaps 8;

Qy 892 GKGIOLKARAFELSEBGVLYKYDGNNSARLLGVFSHSEWAGYTLNSTSSSLMDFTTDA 951
Db 15 GKRIQLARAFELSEBGVLYKYDGNNSARLLGVFSHSEWAGYTLNSTSSSLMDFTTDA 74
Qy 952 ENTSSKGFELHFSFELIKCEDPGTPKRYKVHDEGHPAGSSVSFSCDPPGYSLRGSEELL 1011
Db 75 ENTSSKGFELHFSFELIKCEDPGTPKRYKVHDEGHPAGSSVSFSCDPPGYSLRGSEELL 134

```

QY 1012 LSGERTWDRPLPTCYAECGGTVRGVSGOVLSPGPAPYENHNCITWTIEACCTIGL 1071
DB 135 LSGERRTWRDRPLPTCYAECGGTVRGVSGOVLSPGPAPYENHNCITWTIEACCTIGL 194
QY 1072 HELVFTDEEVDLRIWMDGVESGVLKLELSPALPKDLSTFNSVLIQFSTDFTSKOG 1131
DB 195 HFLVFTDEEVDLRIWMDGVESGVLKLELSPALPKDLSTFNSVLIQFSTDFTSKOG 254
QY 1132 FALQFVSTATSCNDGIPONGSRGSDSWEADSTVFCODPGYALQGSALISCVKIENRF 1191
DB 255 FALQFVSTATSCNDGIPONGSRGSDSWEADSTVFCODPGYALQGSALISCVKIENRF 314
QY 1192 FMOQSPPTCIAPCGGDLTGPBGVILSPNYPPEYPCKECDMKVTYSPDYVALVFNFL 1231
DB 315 FMOQSPPTCIAPCGGDLTGPBGVILSPNYPPEYPCKECDMKVTYSPDYVALVFNFL 374
QY 1252 BEGYDPLHLYDGRDLSPLIGSFYSGOLPGR1ESSNSLFLAFRSDASVSNAGFYIDYTE 1311
DB 375 BEGYDPLHLYDGRDLSPLIGSFYSGOLPGR1ESSNSLFLAFRSDASVSNAGFYIDYTE 434
QY 1312 NPRESCEPDGSIKNGTRVSDKLGSSVTVYCHGGEVEGTSCLIGPDGKVMNNPR 1371
DB 435 NPRESCEPDGSIKNGTRVSDKLGSSVTVYCHGGEVEGTSCLIGPDGKVMNNPR 494
QY 1372 PICTAPCGGQYVGSQVILSPNYPONTSGOICLYFTVTPKDYVVFQGFALNDVY 1431
DB 495 PICTAPCGGQYVGSQVILSPNYPONTSGOICLYFTVTPKDYVVFQGFALNDVY 554
QY 1432 EVDHDSHSHSLSLSGSHT----- 1452
DB 555 EVDHDSHSHSLSLSGSHT----- 1452
QY 1453 ----- 1452
DB 615 CLSCGACTQRSQLVRAPTSGAFSCPHPCVYTAPLMCSLLLNQNTNMLQVULVS 674
QY 1453 ----- 1452
DB 675 LPWPLCTAPSRRTYFVFCYKSCOSTLVSCAHAGSLPLATSNQVLIKFSKAGLAPARGFI 734
QY 1481 FVYQ-----AVPRTSATQSSVPEPRYKRLGSPFSGALVREPCNSGVYALQSPFI 1532
DB 735 FVYQMEDMDAGVPRISATQSSVPEPRYKRLGSPFSGALVREPCNSGVYALQSPFI 794
QY 1533 ECLFVPBALQOMNVSATPCVPCCGNLTERRGITLSPGPEPYLNSLNCVWKIVPEGAG 1592
DB 795 ECLFVPBALQOMNVSATPCVPCCGNLTERRGITLSPGPEPYLNSLNCVWKIVPEGAG 854
QY 1593 IOIOVSEFVTEQNMDSLEVPDAGDNTVTLGSGFGTTVPALLNSTNOLYHFRYSISVS 1652
DB 855 IOIOVSEFVTEQNMDSLEVPDAGDNTVTLGSGFGTTVPALLNSTNOLYHFRYSISVS 914
QY 1653 AAGHLEKTVGLSSCEPAVPSNGVYTGERYLVNDVVSFOCEPGYALQGHASISCMPT 1712
DB 915 AAGHLEKTVGLSSCEPAVPSNGVYTGERYLVNDVVSFOCEPGYALQGHASISCMPT 974
QY 1713 VRRANYPPTCIAQCGGTVEMEGVILSPFPENYNSMDCSMKIALPVGFAGHIOPLNF 1772
DB 975 VRRANYPPTCIAQCGGTVEMEGVILSPFPENYNSMDCSMKIALPVGFAGHIOPLNF 1034
QY 1773 STEPNHDIYIRNGPYETSRMGRFSGSELPSLSTSHETTVYFHSHDSQRKPFKLEY 1832
DB 1035 STEPNHDIYIRNGPYETSRMGRFSGSELPSLSTSHETTVYFHSHDSQRKPFKLEY 1094
QY 1833 Q-----AYELOCEPDPEPFANGI 1850
DB 1095 QDLTYHQJSSFLRGDLSELRINSTPVAASYWDLDPGCEAYELOCEPDPEPFANGI 1154
QY 1851 VRGAGYVNGSVYFECLEPGYOLTGHPVLTQCGHTRNMDHPLPKCEVCCGNTITSNTV 1910
DB 1155 VRGAGYVNGSVYFECLEPGYOLTGHPVLTQCGHTRNMDHPLPKCEVCCGNTITSNTV 1214

QY 1911 YSPGFPSSSODCWLITV-PIGHVRLNLSLQTERPSGPRITIMDPOQOTAPRLGVF 1969
DB 1215 YSPGFPSSSODCWLITVQNLANGVRLNLSLQTERPSGPRITIMDPOQOTAPRLGVF 1274
QY 1970 TRSMAKTVOSSNOVLKFRHDATGFIAPAFSAVPLTKCPPTILPNAEVTENEBF 2029
DB 1275 TRSMAKTVOSSNOVLKFRHDATGFIAPAFSAVPLTKCPPTILPNAEVTENEBF 1334
QY 2030 NIGDIYRCLPFTLVNGBEILTKLGTLOJEGPPICEVHCPTNELLTDSGTVILSOS 2089
DB 1335 NIGDIYRCLPFTLVNGBEILTKLGTLOJEGPPICEVHCPTNELLTDSGTVILSOS 1394
QY 2090 YPGSYQFOTCSMLVREPDYNI SLTVEYFLSEKQYDEFEIPDGSGOPLKALSGVNS 2149
DB 1395 YPGSYQFOTCSMLVREPDYNI SLTVEYFLSEKQYDEFEIPDGSGOPLKALSGVNS 1454
QY 2150 APLIYVSSNSVYLRSSDHAHNRGPFIRYAPCSLPRADLHGFILQSTOPGGSIH 2209
DB 1455 APLIYVSSNSVYLRSSDHAHNRGPFIRYAPCSLPRADLHGFILQSTOPGGSIH 1498
QY 2210 FCGNAGYRLVGHSMALCTRHPOGYHLMSEALPLCOALSCLPPEARKNMGVFGKETTGTK 2269
DB 1499 FCGNAGYRLVGHSMALCTRHPOGYHLMSEALPLCOALSCLPPEARKNMGVFGKETTGTK 1558
QY 2270 AVYSGSEGYHLOAGAEATAECLDTGLMSNRVPPQCV-----PYTCPDVSSISYEH 2320
DB 1559 AVYSGSEGYHLOAGAEATAECLDTGLMSNRVPPQCV-----PYTCPDVSSISYEH 1618
QY 2321 GRWRLIFETOYQFOQMLMICDPGYYYTGORVIRCOANAKXSLDSTPTCCIIISCGELPI 2380
DB 1619 GRWRLIFETOYQFOQMLMICDPGYYYTGORVIRCOANAKXSLDSTPTCCIIISCGELPI 1678
QY 2381 PENGRIRITLSYGGATAFSCNSGYTLVGSVRECMANGLSGSEVRCL----- 2429
DB 1679 PENGRIRITLSYGGATAFSCNSGYTLVGSVRECMANGLSGSEVRCL----- 1738
QY 2430 -----AGHCGTEPIVNGHINENSVYSGVYVQCNAPFRLIGMSVRLCOQ 2475
DB 1739 KULFVULSSPSLTYAGHCGTEPIVNGHINENSVYSGVYVQCNAPFRLIGMSVRLCOQ 1798
QY 2476 DHMSGKTPFCV-----PITGHPGNVNLJTOGNQPLN 2510
DB 1799 DHMSGKTPFCVHQOQLLLLLLCCDDDDDEDDGSAICGHPGNVNLJTOGNQPLN 1858
QY 2511 DVVFVVCNPGYMAAGARSOCCLASGOWSMLPFCRIINCTDPEGHQENSVPQVHAGPHRF 2570
DB 1859 DVVFVVCNPGYMAAGARSOCCLASGOWSMLPFCRIINCTDPEGHQENSVPQVHAGPHRF 1918
QY 2571 SFGTTVSYRCNHGFYLLGTIVLSQCGDGTWDRRPOCL 2608
DB 1919 SFGTTVSYRCNHGFYLLGTIVLSQCGDGTWDRRPOCL 1956

RESULT 8
US-10-276-934-13
; Sequence 13, Application US/10276934
; Publication No. US20030180750A1
; GENERAL INFORMATION:
; APPLICANT: University of Leeds
; APPLICANT: Markham, Alexander F.
; APPLICANT: Jackson, Andrew P.
; APPLICANT: Woods, Christopher G.
; TITLE OF INVENTION: Treatment of Cancer and Neurological Diseases
; FILE REFERENCE: 9052-144
; CURRENT APPLICATION NUMBER: US/10/276,934
; PRIOR FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: PCT/GB01/02240
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: GB0012186.3
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13

```

LENGTH: 2352
TYPE: PR
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (11)..(11)
OTHER INFORMATION: "x" is unknown amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (212)..(212)
OTHER INFORMATION: "x" is unknown amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (594)..(594)
OTHER INFORMATION: "x" is unknown amino acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (731)..(731)
OTHER INFORMATION: "x" is unknown amino acid
US-10-276-934-13

Query Match 45.1%; Score 7662.5; DB 14; Length 2352;
Best Local Similarity 65.7%; Pred. No. 0;

Matches 1357; Conservative 291; Mismatches 341; Indels 77; Gaps 11;

1 MAGAPPAL-----LPCSLISD-----CCASNRHSVSGVPSSELVKQIELKRGV 47
310 ISGEADPSIMLTGNLSPVSSKNWLRHFTSDSNHRR-GFNAQFOVKALIELKRGV 368
48 KLMPKSNOKTSVLTGVVSGHNMCPDPCI PERGKRLSGDFRLGSSVQFTCEGYDLQ 107
369 KMLPSKSHKNSVLSOGVALVSDMCPDPCI PENGRRAGSDFRVGNANVQSCSDNYLQ 428
108 GSKKITCKKSDMAANSRHRPVCRAMCDALRGPSGIIITSVPPIQYDNNACWIIIT 167
429 GSKKITCKKSDMAANSRHRPVCRAMCDALRGPSGIIITSVPPIQYDNNACWIIIT 488
168 ALNRSKVIKLAFFEEFDLERYDVLTVGGGODGOKTVLMSQACSDPHTPGSRIPR- 226
489 TTDDKVIKLAFFEEFDLERYDVLTVGGGODGOKTVLMSQACSDPHTPGSRIPR- 538
227 --SNSGDIWROKMTVLICRDISSDARSQSVKSPKTSNAVELVAPG---TEIRGS 279
539 IVSMNQMLH-----LQSDDS-----IGSPGFAVVOEIRKGS 572
280 CGDPGICPAYGRRESRRHNGDTLKFECOPAFELVQKAITCQKNQMSAKKPGCVFSCFP 339
573 CGDPGICPAYGRRESRRHNGDTLKFECOPAFELVQKAITCQKNQMSAKKPGCVFSCFP 632
340 NFPSPGVVLSPNYPEDYGNHLHCVLILAPESRIHLAFNDIVEPOQDFLVIKIDATA 399
633 NFPSPGVVLSPNYPEDYGNHLHCVLILAPESRIHLAFNDIVEPOQDFLVIKIDATA 692
400 EAPVLTGFSGNQLPSSITSSGVARLEFQTHSTGKGFNITFTFRHNECPDGVVNG 459
693 DITVLGTFSGNEVSQLASSGHIVRLERQSHSTTGKXNITTYTTCGNECHDGIPLNG 752
460 KRPEDSLQSSISFLCDEGFLTGQSETTICVLKEGSVVNSAVLRCEAPCGHILSPS 519
753 RRFEDRELGSVFFHCDGDFVKTQGSRSITCIIQDGNVWVSSTVPRCEAPCGHILSPS 812
520 GTILSPMPGPKYKALSCAWIEAOGPYKITEDRFTEVNYVLTLEPRORTYSAPLIG 579
813 GVILSPMPGPKYKALSCAWIEAOGPYKITEDRFTEVNYVLTLEPRORTYSAPLIG 872
580 VYHGTQVQPLISTSNLYLLFTDQSHSDIGFOLRYETITLQSDHCLDPIPVNGORHG 639
873 EYHGTQVQPLISTSNLYLLFTDQSHSDIGFOLRYETITLQSDHCLDPIPVNGORHG 932
640 NDFVYVGLVTRSCSGITLSDGEPLECPNFQMSRALPSCALCGGFIQSSGTTLSRPF 699
933 GDFGIRSTVTFSCDPGYTLSDDEPLVCERNHQMHALPSCALCGGYIQGSGTIVLSRPF 992

QY 700 PDFYNNLNTCTWIIETSHGKGVFTFTFTHLESCHDYLITENSGFTQPLROLTGRSLPA 759
DB 993 PDFYNNLNTCTWIIETSHGKGVFTFTFTHLESCHDYLITENSGFTQPLROLTGRSLPA 1052
QY 760 PISAGLYGNFTAOVRFISDFSMSEGNITFSEYDLEPCEPEVPAVSIRKGLQFGVGT 819
DB 1053 TIRAGLFGNFTAOVRFISDFSMSEGNITFSEYDLEPCEPEVPAVSIRKGLQFGVGT 1112
QY 820 LTFSCFPGYVLEGRARTTCGRRRLMSSPLPCVACGNSVYCTOCTLLSPNPNVYN 879
DB 1113 LTFSCFPGYVLEGRARTTCGRRRLMSSPLPCVACGNSVYCTOCTLLSPNPNVYN 1172
QY 880 NHECTIYIOTOPGKGIOLKARAFELSGDYLVKYDGNMNSARLLGVSHSEMMGVNTNST 939
DB 1173 NHECTIYIOTOPGKGIOLKARAFELSGDYLVKYDGNMNSARLLGVSHSEMMGVNTNST 1232
QY 940 SSSIMLDPITDAENTSKGFLHFSFELICEDPGTPKFGYKVHDEGHFAGSSVSFSCDP 999
DB 1233 SNHLMLEFNTNGSDTDGQFQTVTSFDLVKCEBPGIYNVGYRIRDEGHFTDVLVYSCNP 1292
QY 1000 GYSLRGEELLCSGERRTWDRPLPTCVACCGTVRGEVGOVLSRGPYAPYENHNCIW 1059
DB 1293 GYAMHGNTLTCISGRVWDKPLPSCIAACGQIHAATSGRILSPGYAPAYDNNLCTW 1352
QY 1060 TIEACCTTGLHFLVPTDEEVHDLRIMDGPVSGVLKELSGPALPKDLHSTFNSVL 1119
DB 1353 TIEADPGKTIHLFIYVDTMADHLKVMGPVDSOILLKEMSGSLRPEIHSTFNSVL 1412
QY 1120 QFSTDFTSKQGAIPQSVSTATSCNDPGIPONGSRSGDSWEAGDSTVFOCDPGYALQS 1179
DB 1413 QFSDPFISKSGFSIQSTSIATCNDPGMPQNGTRVGDREAGDTVTFOCDPGYALQS 1472
QY 1180 AETSCKIENRFPQSPPTCIAACGGDLRGPSGVILSPYPRYPRKCKDMKVVYSPD 1239
DB 1473 AKITCVQNLNRFPMQDPPTCIAACGGDLRGPSGVILSPYPRYPRKCKDMKVVYSPD 1532
QY 1240 YVIALVFNILNLEBGVFLHIYDGRDLSPLIGSFYSGQIPGRJESSNSLFLAFRSDAS 1299
DB 1533 YVIALVFNILNLEBGVFLHIYDGRDLSPLIGSFYSGQIPGRJESSNSLFLAFRSDAS 1592
QY 1300 VSNAGFVIDYTERPESCFDPGSIKNGTRVSGDLKLGSSVYVYCHGGEVEGTSTLSCIL 1359
DB 1593 VGLSGFAIFBKEXKREACFPDGNIMNGTRVGTDFKLGSTITTYOQDSGKYLIDPSSITCVI 1652
QY 1360 GPDKPVMNRPVPCIAACGGGVVSGVLSRPNYPONTYSGQICLYFVVPVDPVYVQ 1419
DB 1653 GADGKPSMDVLPSCNAPCGGYTGSBYVLSPPYPNYTAGOICLYSIVPEKFFVFGQ 1712
QY 1420 FAFPHLNDVVEVHDGSHSRLLSLSGSHGESLPLATSNQVLKIFSAKGLAPARGF 1479
DB 1713 FAFPHLNDVVEVHDGSHSRLLSLSGSHGESLPLATSNQVLKIFSAKGLAPARGF 1772
QY 1480 HFVYQAVPRTSATQCSVPEPRYKRLGSDFSVGAIVRPECNSGVALQGSPEIECLVP 1539
DB 1773 HFVYQAVPRTSATQCSVPEPRYKRLGSDFSVGAIVRPECNSGVALQGSPEIECLVP 1832
QY 1540 ALAQMNVSAATCVVPCCGNLTBRGTLISGFPPEPYINSLNACWKIYVPGAGIQLQVVS 1599
DB 1833 ALAQMNVSAATCVVPCCGNLTBRGTLISGFPPEPYINSLNACWKIYVPGAGIQLQVVS 1892
QY 1600 FVTEQNMDSLEVPDGAINTVLMGSPGTVLPALNLSNOLYHFFYSDISVSAAGFHL 1659
DB 1893 FVTEQNMDSLEVPDGAINTVLMGSPGTVLPALNLSNOLYHFFYSDISVSAAGFHL 1952
QY 1660 YKTVGLSSCEPEAVPSNGVTRGRIYVNDVVSFOCEBGVALQGHAIISMPGVRRVNP 1719
DB 1953 YKTVGLSSCEPEAVPSNGVTRGRIYVNDVVSFOCEBGVALQGHAIISMPGVRRVNP 2012
QY 1720 PPLCIAOCGTVBEMEVILSPGPGYVPSNMCMKIALPVGRGAIIOFLNFTSTERNH 1779
DB 2013 PPLCIAOCGTVBEMEVILSPGPGYVPSNMCMKIALPVGRGAIIOFLNFTSTERNH 2072
QY 1780 YIEIRNPFYETSRMGRFSGSELPSLLSTSHETTVYFHSDHQNRPGFKLEYQAYELQE 1839

```

Db      2073 FLEIQNGPYHTSMIGOFSGTDLPAALISTHETLHFFSDHSQNRQKLAIVQAYELQON 2132
Qy      1840 CPDPEFANGIVAGAAVYNGQSTFECLEGYOLTGHPVLTCHGNRMNDHPLPCCEVPC 1899
Db      2133 CPDPEFANGIVAGAAVYNGQSTFECLEGYOLTGHPVLTCHGNRMNDHPLPCCEVPC 2192
Qy      1900 GGNITSSNGTVVSPGFPSPYSSQDCWMLITVPIGHVRLNLSLQTEPSGDFTIMDGP 1959
Db      2193 GYVVTSONGTIVSPGFPDEYPLIKOCIMLITVPHGVYINFTLQTEVANDYIAVWDGP 2252
Qy      1960 QGTAPRLGVFTSMMAKVTYQSSNOYLLKFRDAATGGIFALAFSAY-----PLTK----- 2010
Db      2253 DQNSPOLGVFSGNTALETVASSSTNOYLLKFRHSDFSNGGFPVLPFHGOLITPTLVKTEMSM 2312
Qy      2011 -----CPPP-----TILPNAEVVTEN 2026
Db      2313 WCLLOCCPTPCFQKFLDSAEGVDS 2338

```

```

RESULT 9
US-10-276-934-14
; Sequence 14, Application US/10276934
; Publication No. US20030180750A1
; GENERAL INFORMATION:
; APPLICANT: University of Leeds
; APPLICANT: Markham, Alexander F.
; APPLICANT: Jackson, Andrew P.
; APPLICANT: Woods, Christopher G.
; TITLE OF INVENTION: Treatment of Cancer and Neurological Diseases
; FILE REFERENCE: 9052-144
; CURRENT APPLICATION NUMBER: US/10/276,934
; PRIOR FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: PCT/GB01/02240
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: GB0012186.3
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 2306
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (166)..(166)
; OTHER INFORMATION: "X" is unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (454)..(454)
; OTHER INFORMATION: "X" is unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (482)..(482)
; OTHER INFORMATION: "X" is unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (548)..(548)
; OTHER INFORMATION: "X" is unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (685)..(685)
; OTHER INFORMATION: "X" is unknown amino acid
; NAME/KEY: MISC FEATURE
; LOCATION: (1013)..(1013)
; OTHER INFORMATION: "X" is unknown amino acid
US-10-276-934-14

```

Query Match 45.0%; Score 7643.5; DB 14; Length 2306;
 Best Local Similarity 65.6%; Pred. No. 0;
 Matches 1355; Conservative 290; Mismatches 344; Indels 77; Gaps 11;

```

Qy      1 MAGAPPALL-----LPCSLISD-----CCASNRHSVGVGPELVKQIELKRGV 47
Db      264 ISGTETASPIWLTGMNLPSPVSISSKNWRLHFTSDSHRRK-GRNAFOYKKAIELSRGV 322
Qy      48 KLMPKDNOKSTVULVOVQSGHNNCPDPGIPERKRLGSDPRLLSSVOFTCNESYDIQ 107
Db      323 KLMPKDNOKSTVULVOVQSGHNNCPDPGIPERKRLGSDPRLLSSVOFTCNESYDIQ 382
Qy      108 GSKRITCMKVSDMFAAMSDHRRPVCARMCDALHRCGSGIITSBNPIQYDNNACWIIIT 167
Db      383 GSKRITCMKVSDMFAAMSDHRRPVCARMCDALHRCGSGIITSBNPIQYDNNACWIIIT 442
Qy      168 ALMPKSVIKLAEFEPLLEGYDITLVGDCGQDQDQKTVLYMSQNAQSDSPHTPGSRIP- 226
Db      443 ALMPKSVIKLAEFEPLLEGYDITLVGDCGQDQDQKTVLYMSQNAQSDSPHTPGSRIP- 492
Qy      227 --SMGDIWRQKWTVLICRDISSPARAGSRKSPKSNAAVELVAPG-----TIEQGS 279
Db      493 IYMSKNQMLH-----LQSDDS-----IGSPGKAVYQIEKGG 526
Qy      280 CGDPGIPAYGRBEGSRFHHGDTLKECPAPFELVGOKAITCQKNQWMAKRGCVFSCFP 339
Db      527 CGDPGIPAYGRBEGSRFHHGDTLKECPAPFELVGOKAITCQKNQWMAKRGCVFSCFP 586
Qy      340 NFTSPGCVVLSPNYEDYDGNHLCWLLIARPRESIHLAFNDIDVEPODFLVIQDQATA 399
Db      587 NFTSPGCVVLSPNYEDYDGNHLCWLLIARPRESIHLAFNDIDVEPODFLVIQDQATA 646
Qy      400 EAPVLGTEPSGNLPSITSSGHVARLEFQDHTSTKRGFNITFTFRHNECCDPVPVNG 459
Db      647 DITVLTGTFSGNEVFPQLASSGHIVALEFQDHTSTKRGFNITFTFRHNECCDPVPVNG 706
Qy      460 KRFQSLQGSISLCEBGLTQGSSTITCVLEKESVWNSAVLRCEAPCGHLLTSPS 519
Db      707 KRFQSLQGSISLCEBGLTQGSSTITCVLEKESVWNSAVLRCEAPCGHLLTSPS 766
Qy      520 GTILSPGMPGFYKDALSCAWYEAOPGVYKITDRPFTEVNYDTLVRDERTYSAPLIG 579
Db      767 GTILSPGMPGFYKDALSCAWYEAOPGVYKITDRPFTEVNYDTLVRDERTYSAPLIG 826
Qy      580 VYHGTQVQPLISTSNVLYLLPSTDKSHSDIGFQRLRYETITLQSDHCLDPGIPVNGORHG 639
Db      827 EYHGTQVQPLISTSNVLYLLPSTDKSHSDIGFQRLRYETITLQSDHCLDPGIPVNGORHG 886
Qy      640 NDFYVGLVTFSCDQSGVTLSDGEPLCEPNQWRALPSCGALCGGFIQSSGTTILSPGF 699
Db      887 GDFGRLVTFSCDQSGVTLSDGEPLCEPNQWRALPSCGALCGGFIQSSGTTILSPGF 946
Qy      700 PDFYVNNINCTWIIETSHGKGVFTFTPHLESCHDYLLITENGSPFOPLROLTGSRLPA 759
Db      947 PDFYVNNINCTWIIETSHGKGVFTFTPHLESCHDYLLITENGSPFOPLROLTGSRLPA 1006
Qy      760 PISAGLVNFTPAQVRFISDPSMSGEGNITFSEYDLEPCEPEVPAVSIRKGLQFGVGD 819
Db      1007 TIKAGLVNFTPAQVRFISDPSMSGEGNITFSEYDLEPCEPEVPAVSIRKGLQFGVGD 1066
Qy      820 LTFSCFPYRLGEGTARITTCGGRRLWSSPLPCVNAECGNSVTGCTGLSPNFPVNNYN 879
Db      1067 LTFSCFPYRLGEGTARITTCGGRRLWSSPLPCVNAECGNSVTGCTGLSPNFPVNNYN 1126
Qy      880 NHECIYSIQTPQKGIQKARAFELSGDVLKYDGNNSARLLGVFSHSEMMAGVTLNST 939
Db      1127 NHECIYSIQTPQKGIQKARAFELSGDVLKYDGNNSARLLGVFSHSEMMAGVTLNST 1186
Qy      940 SSSLMLDFITDAENTSGPELHSSPELLKCEBPGRPKGKYKVDGCHRAAGSVSFCDP 999
Db      1187 SSSLMLDFITDAENTSGPELHSSPELLKCEBPGRPKGKYKVDGCHRAAGSVSFCDP 1246
Qy      1000 GYSLRGSEELLCLSGERRTWDRPLPTCVACCGGTVAEGSVGVLYSGYVAPAYEHNINCM 1059
Db      1247 GYSLRGSEELLCLSGERRTWDRPLPTCVACCGGTVAEGSVGVLYSGYVAPAYEHNINCM 1306
Qy      1060 TIEAEACCTIGLHFLVDTBEVHDVLRIMDGPVESGVLLKELSGPALPKDLHSTFNSVVL 1119

```



```

1307 IIEADPGTISLHFIIVFOTEMAHDILKWMGDVPDSDILKXMGSSALPEDIHSTFNSITL 1366
1310 QFSDPFTSKGCFIORSVSTATSCNDPITPONSRSQSDSHEADSTVFOCDPGYALOGS 1179
1367 QFSDPFTSKGCFIORSVSTATSCNDPITPONSRSQSDSHEADSTVFOCDPGYALOGS 1426
1180 AECVKEIENFEPFOPSPPTCIAPCGGDLTGPSCGVLSPNPEPPPEKCDMWKVTSPD 1239
1427 AKITCVQJNNRFFQOPDPPTCIACCGNLTGPAGVILSPNPEPPPEKCDMWKVTSPD 1486
1240 VYIALVFNIFNLBEGYDFLHITDGRDLSPLIGSFYGSQDLPGRIEISSNSLFLAFRSDAS 1239
1487 FVIALIFKSPMEBSYDFLHIEGSDNSPLIGSFYGSQDAERIESSNSLFLAFRSDAS 1546
1300 VSNAQFVIDYENRESCEPDGSIKNGTRVGSDDLGSSTVYCHGIEVEGISTLSCITL 1359
1547 VGLSGFAIEFKKREACFDPGNIMNGTRVGTDFLGSSTVYCHGIEVEGISTLSCITL 1606
1360 GPDGKPVWNNRPFVCTAPCGGQYVGSQDVLSPNPEPPPEKCDMWKVTSPD 1419
1607 GADKPSMDVLPSCNAPCGQYVGSQDVLSPNPEPPPEKCDMWKVTSPD 1666
1420 FAFHTALNDVNEVHDGSHQSRLLSLSGSHTESSLPATSNQVLKFSKAGLAPARGF 1479
1667 FAYFQTLNDLAELEFDGTHAQARLLSLSGSHTESSLPATSNQVLKFSKAGLAPARGF 1726
1480 HFVYQAVPRTSATOCSSVPEPRYGRKGLSDPSVGAIVRECNQSGYALOGSEPIECLPYPG 1539
1727 HFVYQAVPRTSATOCSSVPEPRYGRKGLSDPSVGAIVRECNQSGYALOGSEPIECLPYPG 1786
1540 ALAOMNVSAPTCVPCCGNLTGERGTLSPGPEPRYGRKGLSDPSVGAIVRECNQSGYALOGSEPIECLPYPG 1539
1787 ALAOMNVSAPTCVPCCGNLTGERGTLSPGPEPRYGRKGLSDPSVGAIVRECNQSGYALOGSEPIECLPYPG 1846
1600 FVTEQNDLSLEVPDADNTVTMLGSFSGTTPALNLTSTNQLYHFFYSDISVSAAAGFLE 1659
1447 FATEQNDLSLEVPDADNTVTMLGSFSGTTPALNLTSTNQLYHFFYSDISVSAAAGFLE 1906
1660 YKTVGLSSCPPEPAPVPSNGVKTGERYLVNDVVSFOCEPEYALOGIAHISCMFETVRNNYP 1719
1907 YKTVGLSSCPPEPAPVPSNGVKTGERYLVNDVVSFOCEPEYALOGIAHISCMFETVRNNYP 1966
1720 PPLCIACCGTVEBMEGVILSPGPGNVPNSMDCSKTALPVGGAHIOPLNFTSTERNHD 1779
1967 SPLCIACCGTVEBMEGVILSPGPGNVPNSMDCSKTALPVGGAHIOPLNFTSTERNHD 2026
1780 YIEIRNGPYETSRMNGRFSGSELSLSTSHETTVYFHSQNSQNRPGFKLEYOAYELOE 1839
2027 FLEIQNGPYHNSPMIGQSGSDTLPAALLSTHETTLHIFYSQNSQNRPGFKLEYOAYELOE 2086
1840 CPDPEPPANGIVRGAGVNGOSVTFECLPGYQLTGHVPLVTCOHGTNRNMDHPLKCEVPC 1899
2087 CPDPEPPANGIVRGAGVNGOSVTFECLPGYQLTGHVPLVTCOHGTNRNMDHPLKCEVPC 2146
1900 GGNITSSNGTVYSGPSPSSQDCWLVLPVIGHGVRMLSLLOTEPSSDFTITINDGP 1959
2147 GYNTSSNGTVYSGPSPSSQDCWLVLPVIGHGVRMLSLLOTEPSSDFTITINDGP 2206
1960 QQTAPRLGVFTSRNAKKTVOSSNOVLKPHRDATGIFAIASAY----PLTK----- 2010
2207 QQTAPRLGVFTSRNAKKTVOSSNOVLKPHRDATGIFAIASAY----PLTK----- 2266
2011 -----CPP-----TILNPAEVTEN 2026
2267 WCLIQCCPTPCFOLKFLDSAGVYDS 2292

```

RESULT 10
 US-10-276-934-11
 ; Sequence 11, Application US/10276934
 ; Publication No. US20030180750A1
 ; GENERAL INFORMATION:

```

APPLICANT: University of Leeds
APPLICANT: Markham, Alexander F.
APPLICANT: Jackson, Andrew P.
APPLICANT: Woods, Christopher G.
TITLE OF INVENTION: Treatment of Cancer and Neurological Diseases
FILE REFERENCE: 9052-144
CURRENT APPLICATION NUMBER: US/10/276,934
PRIOR FILING DATE: 2002-11-20
PRIOR APPLICATION NUMBER: PCT/GB01/02240
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: GB0012186.3
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 2008
TYPE: PRT
ORGANISM: Homo sapiens
US-10-276-934-11

Query Match 44.6%; Score 7568.5; DB 14; Length 2008;
Best Local Similarity 66.7%; Pred. No. 0;
Matches 1328; Conservative 285; Mismatches 316; Indels 63; Gaps 8;

62 LTGVGQGHNMCPDPGIPERGRKLGSDFRGSSVFTCEGVDLQSKRTTCKVSDMF 121
39 VSQGVALLVSDMCDPDPGIPERGRKLGSDFRGSSVFTCEGVDLQSKRTTCKVSDMF 98
122 AAWSDHPRVCRARMCDAHLRGPSTIITSNPEPIQYDNNACWIIITLNPBKVIKLAEE 181
99 AAWSDHPRVCRARMCDAHLRGPSTIITSNPEPIQYDNNACWIIITLNPBKVIKLAEE 158
182 FDLERGVDTLVGEGGDDGQKTYLVMSQNAACSHTPSGRIRPE---SNSGDITWRQKT 238
159 FELERGVDTLVGEGGDDGQKTYLVMSQNAACSHTPSGRIRPE---SNSGDITWRQKT 205
239 VLEICRDISSDARSQSVKSPKTSNAVELVAPG-----TEIEGSGCDPPIPAYGRREG 293
206 -----LQSDDS-----IGSPFKAVYQIEKGGCGDPIPAYGRREG 242
294 SRFHAGDTLKECPAPAEVYQKAITQKNNQWSAKKPGCVFSCFNFSTPSGVVLSPNY 353
243 SSFLAGDTLKECPAPAEVYQKAITQKNNQWSAKKPGCVFSCFNFSTPSGVVLSPNY 302
354 PEDYGNHLACWILLAPBESRILHAFNDIDYEPDFVLVIXDQATAPVLTGTSSGNLP 413
303 PEEYGNMNCWMLIISBPSGRILHAFNDIDYEPDFVLVIXDQATAPVLTGTSSGNLP 362
414 SSITSGHVARLEFQTHSGTGKGFNTFTTFRNECPDGPVNGKRGFGDLSQSSIS 473
363 SOLASSGHIVRLEFQTHSGTGKGFNTFTTFRNECPDGPVNGKRGFGDLSQSSIS 422
474 FLCDGELGTQSGSETITCVLKEGSVWNSAVLRCAPCGGHTLSPSGTILSPGMPGEYKD 533
423 FHODDGVKTKQSGSISITCIQDGNVWNSITPREACCGGHTLSPSGTILSPGMPGEYKD 482
534 ALSCAWIIEAPGPIKITDRFKTEVNYDTLEVRDRTYASPLIGYHGTQVQFLIST 593
483 SLHCWIIIEAPGPIKITDRFKTEVNYDTLEVRDRTYASPLIGYHGTQVQFLIST 542
594 SNVYLLFSDKSHSDIGFOLRYETITLQSHGCDLDPGIPVNGGHRGDFVYGVALTPTSCD 653
543 GNFYLLFSDKSHSDIGFOLRYETITLQSHGCDLDPGIPVNGGHRGDFVYGVALTPTSCD 602
654 SGYTLSDGEPLECEPNFQWGRALPSCALCGPIQSGSSGTILSPGPFDPYNNLACTWII 713
603 PGYTLSDGEPLECEPNFQWGRALPSCALCGPIQSGSSGTILSPGPFDPYNNLACTWII 662
714 ETSHGKGVFTFTFTFLHESGHDYLLITENSGFTPOPLKQLTGSRLPAISGLVYNTFAOV 773
663 EVSHGKGVOMI FHTFHLHESGHDYLLITENSGFTPOPLKQLTGSRLPAISGLVYNTFAOV 722
774 RFLSDFSMSYEGFNIITSEYDLCEPCEPVPAYSIRKGLQGVGDVTLTFSGCFPYRLEGT 833

```


423 FHDDDFVKTQSESTICILQDGNVWSSVTPRCBAFCGSHLTAASSGVILPFCMPGYXD 482
534 ALSCANVIAEQEYPIKITDFRKTENVYDTELVDRGRYASAPLIGVHSTQVPOULIST 593
483 SLHCWEIIEAKPKHSIKITFDRFQTEVNYDTLEVRDGPASSSPGISGHOTQAPQELIST 542
594 SNVLYLFSTDKSHSDIGFOLAREYITLLOSDHCLDGPINPGORHNDPVGALVMPESCD 653
543 GNMWYLLFTDNRSSSIGFLIHESVLESCLDPEIPNGHRHGDGFIRSTVTFSCD 602
654 SGYTLSDGEPLCEEPFQWMSRALPSCBACGCFIQGSSGTLSPFPDPFYNNLACTWII 713
603 PGTLSDDEPLVERHNNHMLPSCDLCGVIYIQGSSGTVLSPFPDPFYNNLACTWII 662
714 ETSHGKGVFTFHTPHLESGBDYLITENGSTFQPLKQLTGSRLPAPISAGLYGNFTAY 773
663 EVSHGKGVQMI FHTPHLESGBDYLITENGSTFQPLKQLTGSRLPAPISAGLYGNFTAY 722
774 RFTSDPSMSYEGNITFSEYDLRCEBEPVAYSIRKGLQFQVGDITLTFSCFPYELCT 833
723 RFTSDPSISYEGNITFSEYDLRCEBEPVAYSIRKGLQFQVGDITLTFSCFPYELCT 782
834 ARITCLGRRRLWSSPLPRCAECGNSVTGQTLSPNPNVYNNNHCEITVIGIOTQPK 893
783 TKITCLGGRWVSAIPRCVABEGASVKNEGTILSPNPNYDNNHECITYIETBAGK 842
894 GIOLKARAFELSEGDLVKYDGNNSARLLGVFSHSEMGVTLNSTSSLMDFITDAEN 953
843 GHILRTSRFQLEGGDTLKVYDKDSSRPLGTFTKNELLGLINSTNHLMEFNNGSD 902
954 TSAGFELHSSFEELIKEDPCTKFGKXNDHDEGHFAGSSVSPFCDPYSLRSGEBELCLS 1013
903 TDGFFOLTYTSPFLVKCEDPGIPIYGYRIDEGFTDVTVLVYSCNPNYAHGNTLTCLS 962
1014 GERTRWDRPLPTCAECGTVRGEVSGOVLSPGPAPYENHNCIMTIBACGTLGLHF 1073
963 GDRWDRWDRPLPSCIAECGQIHAATSRILSPGPAPYENHNCIMTIBACGTLGLHF 1022
1074 LVFDEBHDVLRIMDGPVSEGVLLKELSGPALPKDLASTFNSVLOFSTDFTSKOGFA 1133
1023 IVFDEBHDVLRIMDGPVSEGVLLKELSGPALPKDLASTFNSVLOFSTDFTSKOGFA 1082
1134 IOTSVSTATSCNBPICIQNSRSGDSWEAGDSTVFQCDPGVALOGSAEISCVKIEENFEW 1193
1083 IQESTSIAATCNDPQMGQNGTRGDSREAGDTVTFQCDPGVALOGSAEISCVKIEENFEW 1142
1194 QBPPTICIAFCGGLTGPSPGVLISPNYBEPYPRGKEDCMKVYTSPOVIALVFNIFNLEP 1253
1143 QBPPTICIAFCGGLTGPSPGVLISPNYBEPYPRGKEDCMKVYTSPOVIALVFNIFNLEP 1202
1254 GYDFLHLYDGRDLSPLIGSFYGSQOLPGRITESSNSLFLAFRSDASVSNAGFVYDTEPN 1313
1203 SYDFLHLYDGRDLSPLIGSFYGSQOLPGRITESSNSLFLAFRSDASVSNAGFVYDTEPN 1262
1314 RESCIPDPSIKNGTRVCSDLKLGSSVTVYCHGVEVEGTSTLCTILCPDQKPVWNNRPV 1373
1263 REACFPDPSIKNGTRVCSDLKLGSSVTVYCHGVEVEGTSTLCTILCPDQKPVWNNRPV 1322
1374 CTAPCGGVYSGDGVLSPNYPONVYSGOCLVYFVTPKDYVVGCAFHTHLLNDVVEY 1433
1323 CNAPCGGVYSGDGVLSPNYPONVYSGOCLVYFVTPKDYVVGCAFHTHLLNDVVEY 1382
1434 HDGHSQSRLLSLSSGSHTSPLATSNQVLIKFSKAGLAPARGFHVVOAVPRTSATQ 1493
1383 FDGTHAARLLSLSSGSHTSPLATSNQVLIKFSKAGLAPARGFHVVOAVPRTSATQ 1442
1494 CSSVPEPRYGRKLDSTSVGAIVFECNSGYALOGSPBIBCLPYFGALLAOMVNSAPICVY 1553
1443 CSSVPEPRYGRKLDSTSVGAIVFECNSGYALOGSPBIBCLPYFGALLAOMVNSAPICVY 1502
1554 PCGGLNTERGTLISPGFPEPYLNSLNCWKTIVPEAGGIQIOVVSVTGSONDMSLEVPD 1613

1503 PCSGNFTQRCGTLISPGFPEPYLNSLNCWKTIVPEAGGIQIOVVSVTGSONDMSLEIHD 1562
1614 GADNVTVMGSPGCTVPPALNLTSTNQLVLFHFSYSDISVSAAGFHELYKTVGLSSCEPEAV 1673
1563 GGVNAPRAGSPGCTVPPALNLTSTNQLVLFHFSYSDISVSAAGFHELYKTVGLAGCEPEAL 1622
1674 PSNGVATGERVYNDVVSFQCEPGVALQGAHAIISCMFGTRWNNYPPPLICAOCCGTVEE 1733
1623 PENSITIGDRVYNDVVSFQCEPGVALQGAHAIISCMFGTRWNNYPPPLICAOCCGTVEE 1682
1734 MEGVILISPGPPGNYPSNMDCKMIALPVGFAHIOPLNFSTEBNHDYIIRNGPYETSM 1793
1683 IGVILISPGPPGNYPSNMDCKMIALPVGFAHIOPLNFSTEBNHDYIIRNGPYETSM 1742
1794 MGRFSGELPSLSLSTHETTYVYFHSDBONRPGFLVVOAYELOCPDPEPANGIYVG 1853
1743 IGVFSGTDLPAALLSTHETTYVYFHSDBONRPGFLVVOAYELOCPDPEPANGIYVG 1802
1854 AGYVNGSVTFECLPGYQLTGHP 1876
1803 SDYSVQSVTFECLPGYQLTGHP 1825

RESULT 12
US-10-276-934-10
; Sequence 10, Application US/10276934
; Publication No. US20030180750A1
; APPLICANT: University of Leeds
; APPLICANT: Markham, Alexander F.
; APPLICANT: Jackson, Andrew P.
; APPLICANT: Woods, Christopher G.
; TITLE OF INVENTION: Treatment of Cancer and Neurological Diseases
; FILE REFERENCE: 9052-144
; CURRENT APPLICATION NUMBER: US/10/276,934
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: PCT/GB01/02240
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: GB0012186.3
; PRIOR FILING DATE: 2000-05-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1800
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-934-10

Query Match 40.9%; Score 6949; DB 14; Length 1800;
Best Local Similarity 67.9%; Pred. No. 0;
Matches 1216; Conservative 259; Mismatches 272; Indels 44; Gaps 5;

62 LTVQVSGOHNMCPPDPIBERGKRLSGDFLSSVQFTCNEGYDLOGSKRITCMKYSDMF 121
39 VSGGVYALVSDMCPDPIBERGKRLSGDFLSSVQFTCNEGYDLOGSKRITCMKYSDMF 98
122 AANSRHPVRCRAMCDALHGPESGITSPNFIQYNNNAHCWITITLALNPSKIYKIAFEE 161
99 AANSRHPVRCRAMCDALHGPESGITSPNFIQYNNNAHCWITITLALNPSKIYKIAFEE 158
182 FDLERGVDTLVGDGQDQDQKTVLYMSQACDSHTGSRPE---SWSGDIWPKQMT 238
159 FELERGVDTLVGDGQDQDQKTVLYMSQACDSHTGSRPE---SWSGDIWPKQMT 205
239 VLEICRDISSDARSQSVKSPKTSNAVELVABG---TEIQGSCGDPGIDAYGRREG 293
206 -----LQSDDS-----IGSGFAVAVGEIEKGCGDGPIDAYGRREG 242
294 SRFHGDTLKEECQAPAFELVGOALITCOXNNQMSAKKPGCVBSCFNFTSPSGVYLSPNY 353
243 SSFLHGDITLFECPAPAFELVGERVITCOQNNQMSGKPKSCVFCFNFFTASSGIIISPNY 302
354 PEDYGNHLHCWMLIARPSRHLAFNDIDVEPQPFVLVXKQATLEAPVLGTGFSGNQLP 413

```

Db 303 PEEEXNNMNCWMLIISPEGRHILIFNDPVEBPODFLAVKDDGISDITVLGTFSGNEVP 362
Qy 414 SSTTSSGHVAREFQDTHSTGKGFNTFTTFRHNEBPDGVPVNGKRPEDSIQSSIS 473
Db 363 SOLASSSHVAREFQDTHSTGKGFNTFTTFRHNEBPDGVPVNGKRPEDSIQSSIS 422
Qy 474 FLDEGEFLTGSEETTCVLEKESGVNNSAVLRCAPCGHLLTSPSGTILSPMPGPKYK 533
Db 423 FHDDGKVTQGSSESTICILQDGNVWMSVTPRCAPCGHLLTSPSGTILSPMPGPKYK 482
Qy 534 ALSCAWIEAOPGYPIKITFDREKTEVNYDTLEVRDRTYSAPLLGVYHGTQVPOPLIST 593
Db 483 SLHCEWIEAKPGHSIKITFDREKTEVNYDTLEVRDRTYSAPLLGVYHGTQVPOPLIST 542
Qy 594 SNLYLLFSTDKSHSDIGFOLARETITLOSCHLCPDIPNQRHNDPFGVGLVYFSCD 653
Db 543 GNMWYLLFTTNDNRSSISGLIHESVYLESDBCLDPIPNQHRHGDGFGIRSTVTSCD 602
Qy 654 SGYTLSDGEPLCEPNEFOWSRALPSCBALCGGFIQSSGTILSPGPDFFYNNLNCWTIT 713
Db 603 PGYTLSDDEPLVERHNNQWMLALPSCDALCGYIQSSGTIVLSPGPDFFYNNLNCWTIT 662
Qy 714 ETSHGKGVFTFTFTHLESQDYLITENGSTFQPLRQLTGSRLPAPISAGLYGNFTAOY 773
Db 663 EVSHGKGVQWIFHTFHLESQDYLITEDGSFSEBPVARTLGSVLPHTIKAGLFGNFTAOY 722
Qy 774 RFISSDSMSYEGNITFSEYDLPCPEBPVARTLGSVLPHTIKAGLFGNFTAOY 833
Db 723 RFISSDSISIEGNITFSEYDLPCPEBPVARTLGSVLPHTIKAGLFGNFTAOY 782
Qy 834 ARITLCGRRRLWSSPLPRCAACGNSVGTQGTILSPNFPVYNNNHECTIYSIQTOPK 893
Db 783 TKYLCGRRRLWSSPLPRCAACGNSVGTQGTILSPNFPVYNNNHECTIYSIQTOPK 842
Qy 894 GYOLKARAPFLSGDVLKYDGNNSARLLGVSHBMGVTLNSTSSIMLDFITDAEN 953
Db 843 GHILARTSRFQLEFGDYLKYDGNSSRPGLTKELLILINSTSNHMLTFEFTNGSD 902
Qy 954 TSXGFLHFSFELIKCEDPGTKFGYKXVDEGHFPAQSSVSFCSDGYSARGSEELLCS 1013
Db 903 TDGFGYTLTYSFPLVCKEDPGIKNYGRIRDBSHFTDVTVLVLSNCGVAAHSGSNLTLCIS 962
Qy 1014 GERRITDRPLPTCAACGCGTVRGEVSGVLSPPGYAPYEHNLNCITIEAAGCTIGLHF 1073
Db 963 GDRVWMDKPLPSCIAACGCGIHAATSGRLSPGYAPYONNMLHCTMIEADPEKTIStHF 1022
Qy 1074 LVPTTEVNDVLRINGPVESSVLLKELSPALPKULHSTFNSVULOFTSDFTSKQGA 1133
Db 1023 IVDITMAHDILKVMWGPVDSDLILKWSGSLALBEDIHSTFNSLTLQFSDPFISSKGSFS 1082
Qy 1134 IOPSVTATSCNDPGIPONGSRSGDSMEAGDSVTFPCDDPVALQGSABISCVIENRPFM 1193
Db 1083 IOSTSIATCNDPGIPONGSRSGDSMEAGDSVTFPCDDPVALQGSABISCVIENRPFM 1142
Qy 1194 QPSPPTICAPCGGDLTGPAGVILSPNYPEPYRPGKCDMKVTVSPDYIALVFNINLEP 1253
Db 1143 QPDPPTICAPCGGDLTGPAGVILSPNYPEPYRPGKCDMKVTVSPDYIALVFNINLEP 1202
Qy 1254 GYVFLHITDGRDLSPLIGSFYGSQPLGRLESSNSLFLAFRSDASVSNAGPYDYTENP 1313
Db 1203 SYVFLHITDGRDLSPLIGSFYGSQPLGRLESSNSLFLAFRSDASVSNAGPYDYTENP 1262
Qy 1314 RESCPGSIKNGTRVGSDDLKSSVTVYCHGVEYEGTSLSCILGPGKAPVNNPRV 1373
Db 1263 RESCPGSIKNGTRVGSDDLKSSVTVYCHGVEYEGTSLSCILGPGKAPVNNPRV 1322
Qy 1374 CTAPCGGYVSGDVLSPNYPQNTSGQICLYFVTVPKDYVFGGAFPHALNDVAV 1433
Db 1323 CNAPCGGYVSGDVLSPNYPQNTSGQICLYFVTVPKDYVFGGAFPHALNDVAV 1382
Qy 1434 HGHSHSLSLSSSHGESLPLATSNQVLIKESAKGLAPRGHFPYQAVPRTSATQ 1493

```

```

Db 1383 PDGTHAQBRLSLSGSHGETLPLATSNQVLIKESAKGLAPRGHFPYQAVPRTSATQ 1442
Qy 1494 CSVPEPRYGRKLGSDGFSVGAIVRECNAGYALQSSPELECLPYRPAALQWNVASATCV 1553
Db 1443 CSVPEPRYGRKLGSDGFSVGAIVRECNAGYALQSSPELECLPYRPAALQWNVASATCV 1502
Qy 1554 PCGMLTERGCTILSPGPEPLANSNCWKLIVPEGAGIQIQVSVFVEQWMSLEVP 1613
Db 1503 PCGMLTERGCTILSPGPEPLANSNCWKLIVPEGAGIQIQVSVFVEQWMSLEVP 1562
Qy 1614 GADNVTMLGSGFTVPALNLSNOLYHFSQISVSAAGPHLEKYTVGLSSCEPAV 1673
Db 1563 GADNVTMLGSGFTVPALNLSNOLYHFSQISVSAAGPHLEKYTVGLSSCEPAV 1622
Qy 1674 PSNGVTEGRYVNVVSVQCEPGVALQGHAIQSCPGVWVWNPPLICATQCGTVE 1733
Db 1623 PSNGVTEGRYVNVVSVQCEPGVALQGHAIQSCPGVWVWNPPLICATQCGTVE 1682
Qy 1734 MEGVILSPGPNVPSNMDCKWIALPVFGAHIOFLNSTEPNHDYIIRNGPYETSHM 1793
Db 1683 MEGVILSPGPNVPSNMDCKWIALPVFGAHIOFLNSTEPNHDYIIRNGPYETSHM 1742
Qy 1794 MRFSGSELPSLSTSHETTYVPHSDHSONPPEKLEYQAYELCEPDPE 1844
Db 1743 MRFSGSELPSLSTSHETTYVPHSDHSONPPEKLEYQAYELCEPDPE 1793

```

RESULT 13

US-10-276-934-12

/ Sequence 12, Application US/10276934

/ Publication No. US20030180750A1

/ GENERAL INFORMATION:

/ APPLICANT: University of Leeds

/ APPLICANT: Markham, Alexander F.

/ APPLICANT: Jackson, Andrew P.

/ APPLICANT: Woods, Christopher G.

/ TITLE OF INVENTION: Treatment of Cancer and Neurological Diseases

/ FILE REFERENCE: 9052-144

/ CURRENT APPLICATION NUMBER: US/10/276,934

/ PRIOR FILING DATE: 2002-11-20

/ PRIOR APPLICATION NUMBER: PCT/GB01/02240

/ PRIOR FILING DATE: 2001-05-21

/ PRIOR APPLICATION NUMBER: GB0012186.3

/ NUMBER OF SEQ ID NOS: 16

/ SOFTWARE: PatentIn version 3.1

/ SEQ ID NO 12

/ LENGTH: 1783

/ TYPE: PR

/ ORGANISM: Homo sapiens

US-10-276-934-12

Query Match 40.8%; Score 6936; DB 14; Length 1783;

Best Local Similarity 68.1%; Pred. No. 0;

Matches 1213; Conservative 258; Mismatches 266; Indels 44; Gaps 5;

```

Qy 62 LTQVAVSQCNNCPGIPERKGRKLGSDPRLGSSVQFQCNBEGYDQGSKRITCMKYSDF 121
Db 39 VSQGVAVLSDMCPGIPENGRKAGSDPRVAGANVQFSCEDNYVULQGSSTICQRTETL 98
Qy 122 AAMDHRPVCRAKMDAHLRPSGIIITSPNPPIQYDNNNAHCWITITALNPSKVILAPBE 181
Db 99 AAMDHRPVCRAKMDAHLRPSGIIITSPNPPIQYDNNNAHCWITITITDPKVIILAPBE 158
Qy 182 FDLERGDTLTYGCGGQDQKTVLYMSQACSDSPHTPGRIPE---SMGDIYRQKVT 238
Db 159 FELERGDTLTYGAGKGVDRSVLYV-----LTGSSVPLIIVSMNQWMLH--- 205
Qy 229 VLEICRDISSDARSAGSVKSPKTSNAVELVAPG-----TEIQQSGCDPGIPAYGREG 293
Db 206 -----LQSDS-----ISPGKAVAYQIEKGGCDPGIPAYGKRTG 242
Qy 294 SRFHGDITLKECPAPFELVQKAITCOJNNQWMAKKGCVFSCPFNFTSPSGVILSPNY 353

```

243 SSFLHGGTLTFECRAAEVLGERVITCOQNNQNSGNKPCVSCFSEFFFTASSGIIISPNY 302
QY 354 PEDIYGNHLCVWLILAPESRIHLAFNDIDIVEPOFDFVLKDGATAEAPVLGTFSGNOLP 413
Db 303 PEEYGNMNCVWLILISEPGRILHLIFNDPVEPOFDLAVYDDOISDIITVLGTFSGNEVP 362
QY 414 SSTSSGHVAREFQDTHSTGKRGFNITFTFRNNECPDPVNPNGRPFDDLSLOGSSIS 473
Db 363 SQAASSGHIVLEFQSDHSTTGKGFNITFTFGNBECHDPGIPNGRRPDRFLGSSVS 422
QY 474 FLDCDEGLTQOSETTICVLKEGCVWNSAVLRCEAPCGHLTSPSGTIIISPGMPGYKD 533
Db 423 FHCDDEGVKTQGESSEICIIQDGNVWMSVYPRCEACCGHLTASSSVILPFGMPGYKD 482
QY 534 ALSCAWVIEAOPGYPIKITFDREKTEVNYDTLEVDRGRTYSAPLIGVHGTQVPOFLIST 593
Db 483 SLHCEWIIEMAKPGHSIKITFDREKTEVNYDTLEVDRGRTYSAPLIGVHGTQVPOFLIST 542
QY 594 SNLYLILFSTDKSHSDIGPOLRYETITLOSCHCLDPIPNVGNHNDGVYVGAALYMPSSCD 653
Db 543 GNMFTYLLFTDNRSSIGFLIHYESVTLSESDCLDPGIPVNGHNGDGFGRSTVTFSCD 602
QY 654 SGYTLSDGEPLJECENFQWMSRALPSCBACGFGIQQSSGTILSPGPDYFNNLNCWII 713
Db 603 PGYTLSDDEPLVCEERNQMHALPSCALCGGYIQGSGTVLSGPFDFYENSLNCTWTI 662
QY 714 ETSHGKCVFTFHTFHLESCHDYLITENGSTFQPLRLQTLGSRPADISAGLYGNFTAOV 773
Db 663 EVSHGKCVQMIFFHTFHLESCHDYLITENGSTFSEPPVRLTGSVLPHTIKAGLFGNFTAOV 722
QY 774 RFTSDPMSYEGFNITSEYDLERCEBEVPAYVIRKGLQPGVODTLTSPCSPRYLEGT 833
Db 723 RFTSDPMSYEGFNITSEYDLERCEBEVPAYVIRKGLQPGVODTLTSPCSPRYLEGT 782
QY 834 ARTTLCGRRRLWSSPLPRCAVECGNSVTQGTLSPPNPNVNNNHCEIYSIOTOPK 893
Db 783 TXLTCLGGRRWMSAPLRCAVECGNSVTQGTLSPPNPNVNNNHCEIYSIOTOPK 842
QY 894 GIOLKARAFELSGDVLYKVDGNNNSARLLGVPSHSEMMGVTLNSTSSLMDFITDAEN 953
Db 843 GIHLRTSPFLPEBDTLKVDGKXSSSRPLGTFTKNNELLGILINSTNHLMEFNTGSD 902
QY 954 TSHGFEHFSSEFLIKEDDGTGPKFGYKNDCEGFPASSVSFGCDPRYSIARSGEELLCLS 1013
Db 903 TDGSPOLTYTSPDLVKCEDPGIIPYGYRIDEGFTDTVVALYSCNPNYAMHGSYTLCLLS 962
QY 1014 GERRTWDRPLPTCVACCGTVRGEVSGQVLSPGYPARYENHNLNCIWTIEAAGCTIGLHF 1073
Db 963 GDRRWMDKPLPSCIAECCGOIHATSRILISPGYPARYENHNLNCIWTIEAAGCTIGLHF 1022
QY 1074 LVFDTREHVDLRIWDGPVESGVLLKELSGPALPKDLHSTFNSVVLQFSTDFTSKOGFA 1133
Db 1023 IVEFTEMAHDILKVMDDPVDSDILKEMSGSALPEDIHSTFNSVVLQFSTDFTSKOGFA 1082
QY 1134 IORSVSTATCNDPBGITQNSRSGDSMEAGDSTVFQCDPRVALOGSAEISCVKIEKRFEM 1193
Db 1083 IQESTSIATCNDPBGITQNSRSGDSMEAGDSTVFQCDPRVALOGSAEISCVKIEKRFEM 1142
QY 1194 QPSPPTCIACCGGLTGPSGVILSPNYPPEYPRGKECDMKVTVSPDVIALVFNIFLPE 1253
Db 1143 QPSPPTCIACCGGLTGPSGVILSPNYPPEYPRGKECDMKVTVSPDVIALVFNIFLPE 1202
QY 1254 GYDFLHIYDGRDLSPLIGSFYGSQULPGRITSSNSLFLAFRSDASVSNAGFYDYTENP 1313
Db 1203 SYDFLHIYEGEDSNPLIGSFYGSQULPGRITSSNSLFLAFRSDASVSNAGFYDYTENP 1262
QY 1314 RESCFDGSIKNGRVSDDLKLGSSVYTYCHGGEVETSTLSGIIILGPDGKPMANNRPV 1373
Db 1263 REACFPDGNIMNGRVGTDFRGLGTTIYQCDSGKXILDPSSITCVIGADKPSMDQVLP 1322
QY 1374 CTAPCGGQVSGDVVSPVYPOVYTSQGLCTVFTVYKDYVVGQAFPHFALNDVVEV 1433

Db 1323 CNAPCGGQVYSGSEGVVLSPVYPHNYTAGOICLYSTVPEKFEVVGQAFVOTALNDLAE 1382
QY 1434 HDGHSOHSRLSSLSGSHGSESLPLATSNOLIKFSAKGLAPRGHFPVYQAVPRSATQ 1493
Db 1383 FDCGTHAQAARLSSLSGSHGSESLPLATSNOLIKFSAKGLAPRGHFPVYQAVPRSATQ 1442
QY 1494 CSSVPEPRYKGRIGSEPSVAGIIVFPCNSGYALOGSEIECLVPGALQOMNYSAPTCV 1553
Db 1443 CSSVPEPRYKGRIGSEPSVAGIIVFPCNSGYALOGSEIECLVPGALQOMNYSAPTCV 1502
QY 1554 PCGNLTERKRTILSPGPEPYLNSLNCWIKIIVPEAGAGIOLQVSVFVEQNMDSLEVP 1613
Db 1503 PCGNLTERKRTILSPGPEPYLNSLNCWIKIIVPEAGAGIOLQVSVFVEQNMDSLEVP 1562
QY 1614 GADNYTMLGSEGGTVPALNSTSNOLYHFPVSDISVSAAGHLEKTYGLSSCEBPV 1673
Db 1563 GGVTAAPRLGSEGGTVPALNSTSNOLYHFPVSDISVSAAGHLEKTYGLSSCEBPV 1622
QY 1674 PSNGVTERGYLNDVVSFQCEBRYALOGHAHISCMPTVRMNPPLCTIACCGVVEE 1733
Db 1623 PSNGVTERGYLNDVVSFQCEBRYALOGHAHISCMPTVRMNPPLCTIACCGVVEE 1682
QY 1734 MEGVILSPGPGNYPNSMDCSWKIALPVGFAGHIOELNFTSEPNHDIETIRNGPYETSRM 1793
Db 1683 LGGVILSPGPGNYPNSMDCSWKIALPVGFAGHIOELNFTSEPNHDIETIRNGPYETSRM 1742
QY 1794 MGRFSGSELPSSLLSTSHETTVFHSQDHSQNRGFLLEYOA 1834
Db 1743 IGOFSGTDLPPALLSTHETTLHFYSQDHSQNRGFLLEYOA 1783

RESULT 14
US-10-016-248-49
; Sequence 49, Application US/10016248
; Publication No. US20040033491A1
; GENERAL INFORMATION:
; APPLICANT: Alecbrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-218
; CURRENT APPLICATION NUMBER: US/10/016, 248
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/254, 329
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/291, 037
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/255, 648
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 60/297, 173
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/309, 258
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/326, 393
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/315, 639
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 49
; LENGTH: 1048
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-016-248-49

Query Match 24.9%, Score 4226; DB 15; Length 1048;
Best Local Similarity 70.2%; Pred. No. 1,4e-293;
Matches 735; Conservative 153; Mismatches 159; Indels 0; Gaps 0;

QY 718 GKGVFTFTHLESCHDYLITENGSTFQPLRLQTLGSRPADISAGLYGNFTAOVRFIS 777
Db 1 GKGVQMIFFHTFHLESCHDYLITENGSTFSEPPVRLTGSVLPHTIKAGLFGNFTAOVRFIS 60
QY 778 DFMSYEGFNITSEYDLERCEBEVPAYVIRKGLQPGVODTLTSPCSPRYLEGTARIT 837

```

Db 61 DFEISYEGFNITSEYDLEPCDDPGVAFARIRIGPHGVGDSLTFSCFLGYRLGATKLT 120
Qy 838 CUGRRRLWSSPLPRCAECCGNSVYTGCTGTLSPNPNVANNHCECTYSIQTPGKIQI 897
Db 121 CUGGGRWWSAPLRCAECCGNSVYTGCTGTLSPNPNVANNHCECTYSIQTPGKIQI 180
Qy 898 KARAFELSEDDVLYKVDGNNSARLLGVFSEMMGVTLNSTSSIMLDPITDAENTS 957
Db 181 KTRSPQLFEDDTLYKVDGNNSARLLGVFSEMMGVTLNSTSSIMLDPITDAENTS 240
Qy 958 FELHSSFELIKCEDPQTPKFKYVHDEGHFAGSSVVSFCDPGYSLRGSEELLCLSGERR 1017
Db 241 FOLYTSFLDKCEDPGI PMYGYRIRDEGHFTDTVLVYSCNPGYAMHGSYTLTCLSGDRR 300
Qy 1018 TWMRPLPTCAECCGTVRGEVSGVLSFGPAPYENHNLCTIIEADGCTTGLHFLVDP 1077
Db 301 VMDKPLPSCIAECCGQIHAATSGRTILSPGYAPYDNNHCTIIEADGCTTGLHFLVDP 360
Qy 1078 TEEDHVLIRIMDGPVSESVLLKELSGPALPKDLHSTFNSVVLQFTDFFTSKOGFAPIS 1137
Db 361 TEWAHDLKXMDGPVSDILKELSGPALPKDLHSTFNSVVLQFTDFFTSKOGFAPIS 420
Qy 1138 VSTATSCNDPILPONGSRGSDSWAGDSVTFQCDPGYALQSAEISCVKIENRFWQPSR 1197
Db 421 TSAATCNDPDMGPNGRVYDSDREAGDVTFFQCDPGYALQSAEISCVKIENRFWQPSR 480
Qy 1198 PTCIAACGGDLTPRSGVYLSPNPPEPPKCECDMKVTVSPDVIALVFNILFEPGYDF 1257
Db 481 PTCIAACGGDLTPRSGVYLSPNPPEPPKCECDMKVTVSPDVIALVFNILFEPGYDF 540
Qy 1258 LHIYDGRDLSPLIGSFYGSQOLPGRIESSNSLFLAFRSDASVSNAGFVLDVYTEMRESG 1317
Db 541 LHIYDGRDLSPLIGSFYGSQOLPGRIESSNSLFLAFRSDASVSNAGFVLDVYTEMRESG 600
Qy 1318 FDEGSIKNGTRVSGDLKLGSSVTVYCHGYEVEGTSLSCITIGDPKPVNNRPPVCTAP 1377
Db 601 FDEGSIKNGTRVSGDLKLGSSVTVYCHGYEVEGTSLSCITIGDPKPVNNRPPVCTAP 660
Qy 1378 CGGQYVGSQGVLSPPNYPOMYTSQGLVTVVYKPYVVGCAFHTLNDVVEVHDI 1437
Db 661 CGGQYVGSQGVLSPPNYPOMYTSQGLVTVVYKPYVVGCAFHTLNDVVEVHDI 720
Qy 1438 SQSRLLSLSGSHTEGSLPLATSNQVLIKFSAKGLAPAGFHFVYQAVERTSATOCSSY 1497
Db 721 HAARLLSLSGSHTEGSLPLATSNQVLIKFSAKGLAPAGFHFVYQAVERTSATOCSSY 780
Qy 1498 PEBRYGRLGSDFSVGAIVFECNCSGYALQGSPEIECLPVGALAQMNSAPTCVPPCG 1557
Db 781 PEBRYGRLGSDFSVGAIVFECNCSGYALQGSPEIECLPVGALAQMNSAPTCVPPCG 840
Qy 1558 NLTERGCTILSPRPPRYLNSLNCWKIVPBEAGTQIOVVSFVTQNMDSLEVPDADN 1617
Db 841 NLTERGCTILSPRPPRYLNSLNCWKIVPBEAGTQIOVVSFVTQNMDSLEVPDADN 900
Qy 1618 TVTMLGFSSTYVPLALNSTNSQLYLHFYSDIVSAAGFLEKTVGLSGCEPAPVPSNG 1677
Db 901 TARRLGFSSTYVPLALNSTNSQLYLHFYSDIVSAAGFLEKTVGLSGCEPAPVPSNG 960
Qy 1678 VKTGERLYLVNDVVSFOCEPGYALQGHANISCMFGTVRNNYPPPLCIAOCGTVBEMEGY 1737
Db 961 IKIGDRMYVNDVVSFOCEPGYALQGHANISCMFGTVRNNYPPPLCIAOCGTVBEMEGY 1020
Qy 1738 ILSPPGPNYPSNMDCGKWKIALPVGRG 1764
Db 1021 ILSPPGPNYPSNMDCGKWKIALPVGRG 1047

```

RESULT 15
 US-10-016-248-48
 ; Sequence 48; Application US/10016248
 ; Publication No. US20040033491A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aleobrook et al.

```

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-218
; CURRENT APPLICATION NUMBER: US/10/016,248
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/254,329
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/291,037
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/255,648
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 60/297,173
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/309,258
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/326,393
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/315,639
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-016-248-48

Query Match      24.8%  Score 4209; DB 15; Length 1043;
Best Local Similarity 70.3%; Pred. No. 2,36-292;
Matches 731; Conservative 153; Mismatches 156; Indels 0; Gaps 0;

Qy 725 FHTFPLESGHDYLLITENGSTFOTPLRQULGSHLPAPISAGLYGNFTAQVRFISDSMSYE 784
Db 3 FHTFPLESGHDYLLITENGSTFOTPLRQULGSHLPAPISAGLYGNFTAQVRFISDSMSYE 62
Qy 785 GNITFSEYDLPECEPEVPAYSIKGLQFGYGDVLTFTSCFPGYLLEGTARITICGGRR 844
Db 63 GNITFSEYDLPECEPEVPAYSIKGLQFGYGDVLTFTSCFPGYLLEGTARITICGGRR 122
Qy 844 LMSPLPRCAECCGNSVYTGCTGTLSPNPNVANNHCECTYSIQTPGKIQI 904
Db 122 LMSPLPRCAECCGNSVYTGCTGTLSPNPNVANNHCECTYSIQTPGKIQI 182
Qy 905 SEGDLVLYKVDGNNSARLLGVFSEMMGVTLNSTSSIMLDPITDAENTS 964
Db 182 SEGDLVLYKVDGNNSARLLGVFSEMMGVTLNSTSSIMLDPITDAENTS 242
Qy 965 FELIKCEDPQTPKFKYVHDEGHFAGSSVVSFCDPGYSLRGSEELLCLSGERR 1024
Db 242 FELIKCEDPQTPKFKYVHDEGHFAGSSVVSFCDPGYSLRGSEELLCLSGERR 302
Qy 1025 TCVAECGGTVRGEVSGVLSFGPAPYENHNLCTIIEADGCTTGLHFLVDPTEVHDV 1084
Db 302 TCVAECGGTVRGEVSGVLSFGPAPYENHNLCTIIEADGCTTGLHFLVDPTEVHDV 362
Qy 1085 LRIWDPYVESGVLLKELSGPALPKDLHSTFNSVVLQFTDFFTSKOGFAPIS 1144
Db 362 LRIWDPYVESGVLLKELSGPALPKDLHSTFNSVVLQFTDFFTSKOGFAPIS 422
Qy 1145 NDGPILPONGSRGSDSWAGDSVTFQCDPGYALQSAEISCVKIENRFWQPSR 1204
Db 422 NDGPILPONGSRGSDSWAGDSVTFQCDPGYALQSAEISCVKIENRFWQPSR 482
Qy 1205 GGDLLGSPGVLSPPNYPOMYTSQGLVTVVYKPYVVGCAFHTLNDVVEVHDI 1264
Db 482 GGDLLGSPGVLSPPNYPOMYTSQGLVTVVYKPYVVGCAFHTLNDVVEVHDI 542
Qy 1265 DLSPLIGSFYGSQOLPGRIESSNSLFLAFRSDASVSNAGFVLDVYTEMRESG 1324
Db 542 DLSPLIGSFYGSQOLPGRIESSNSLFLAFRSDASVSNAGFVLDVYTEMRESG 602
Qy 1325 NSTRVGSDLKLGSSVTVYCHGYEVEGTSLSCITIGDPKPVNNRPPVCTAPCGGQYVG 1384
Db 602 NSTRVGSDLKLGSSVTVYCHGYEVEGTSLSCITIGDPKPVNNRPPVCTAPCGGQYVG 662

```

[illegible]

Search completed: October 18, 2004, 10:33:18
Job time : 110 secs

This Page Blank (uspio)